

Opinion Dynamics and the Evolution of Social Power in Social Networks

Mengbin Ye



Australian
National
University

A thesis submitted for the degree of
Doctor of Philosophy
The Australian National University

August 2018

© Mengbin Ye 2018
All Rights Reserved.

Except where otherwise indicated, this thesis is my own original work.
The nature and extent of collaboration have been outlined in this thesis.

Mengbin Ye
15 August 2018

To my family.

"A wizard is never late, Frodo Baggins. Nor is he early. He arrives precisely when he means to."

— Gandalf the Grey.

"I wish the ring had never come to me. I wish none of this had happened"

— Frodo Baggins.

"So do all who live to see such times. But that is not for them to decide. All we have to decide is what to do with the time that is given to us."

— Gandalf the Grey.

Declaration

My doctoral studies were conducted under the guidance and supervision of Emeritus Prof. Brian D. O. Anderson, Assoc. Prof. Changbin (Brad) Yu and Assoc. Prof. Jochen Trumpf. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

All of the results in this thesis have been published, or are currently under review, in refereed international journals or international conferences. Some of these results have been achieved in collaboration with other researchers, or carried out in other institutions. A statement is included as part of the dissertation, clearly outlining the extent of collaboration, with whom and under what auspices. Co-authored publications arising during the period of my thesis research are listed as below: (* indicates that the thesis incorporates material of this paper)

Submitted Journal Papers

- [1] ***M. Ye**, M.H. Trinh, Y.-H. Lim, B. D.O. Anderson, C. Yu and H.-S. Ahn. Continuous-Time Opinion Dynamics with Multiple Logically Interdependent Topics. Submitted to *Automatica*
- [2] ***M. Ye**, Y. Qin, A. Govaert, B. D.O. Anderson, and M. Cao. An Influence Network Model to Study Discrepancies in Expressed and Private Opinions. Submitted to *Automatica*.
- [3] J. S. Russell, **M. Ye**, B. D.O. Anderson, H. Hmam and P. Sarunic. Cooperative Localisation of GPS-Denied UAVs using Direction of Arrival Measurements. Submitted to *IEEE Transactions on Aerospace and Electronic Systems*. An extended version is found on ArXiv: <https://arxiv.org/abs/1804.04317>.
- [4] S. Kawamura, K. Cai, **M. Ye**, and Z. Lin. Tight Bound on Parameter of Surplus-Based Averaging Algorithm over Balanced Digraphs. Submitted to *International Journal of Control*.

Refereed Journal Papers

- [5] ***M. Ye**, J. Liu, B. D.O. Anderson, C. Yu and T. Başar. Evolution of Social Power in Social Networks with Dynamic Topology. In *IEEE Transactions on Automatic Control* (regular paper). Online: <https://doi.org/10.1109/TAC.2018.2805261>
- [6] **M. Ye**, B. D.O. Anderson and C. Yu. Leader Tracking of Euler-Lagrange Agents on Directed Switching Networks Using A Model-Independent Algorithm. In *IEEE Transactions on Control of Networked Systems*. Early access: <https://doi.org/10.1109/TCNS.2018.2856298>

- [7] Q. Liu, **M. Ye**, J. Qin and C. Yu. Event-Triggered Algorithms for Leader-Follower Consensus of Networked Euler-Lagrange Agents. In *IEEE Transactions on Systems, Man and Cybernetics: Systems*
<https://doi.org/10.1109/TSMC.2017.2772820>
- [8] **M. Ye**, B. D.O. Anderson and C. Yu. Distributed Model-Independent Consensus of Euler-Lagrange Agents on Directed Networks. *International Journal of Robust and Nonlinear Control*, **27**(14): 2428-2450, September 2017.
- [9] **M. Ye**, B. D.O. Anderson and C. Yu. Bearing-Only Measurement Self-Localization, Velocity Consensus and Formation Control. *IEEE Transactions on Aerospace and Electronic Systems*, **53**(2): 575-586, April 2017.

Refereed Conference Papers

- [10] **M. Ye**, J. Liu and B. D.O. Anderson. On the Effects of Heterogeneous Logical Interdependencies in Multi-Dimensional Opinion Dynamics Models. To appear in *57th IEEE Conference on Decision and Control*, Miami, USA, Dec 2018.
- [11] J. Liu, **M. Ye**, B. D.O. Anderson, Tamer Başar, and Angelia Nedić. Discrete-Time Polar Opinion Dynamics with Heterogeneous Individuals. To appear in *57th IEEE Conference on Decision and Control*, Miami, USA, Dec 2018.
- [12] M. H. Trinh, B.-H. Lee, **M. Ye**, and H.-S. Ahn. Bearing-based Formation Control and Network Localization via Global Orientation Estimation. To appear in *2nd IEEE Conference on Control Technology and Applications*, Copenhagen, Denmark, August 2018.
- [13] *B. D.O. Anderson, and **M. Ye**. Nonlinear Mapping Convergence and Application to Social Networks. To appear in *European Control Conference*, pp. 557-562, Limassol, Cyprus, June 2018.
- [14] ***M. Ye**, J. Liu, B. D.O. Anderson, C. Yu and T. Başar. Modification of Social Dominance in Social Networks by Selective Adjustment of Interpersonal Weights. In the *56th IEEE Conf. on Decision and Control*, pp. 2906-2911, Melbourne, Australia, 2017.
- [15] M.H. Trinh, **M. Ye**, B. D.O. Anderson and H.-S. Ahn. Matrix-Weighted Consensus With Leader-Following Topologies. In *Asian Control Conference*, pp. 1795-1800, Gold Coast, Australia, 2017.
- [16] ***M. Ye**, J. Liu, B. D.O. Anderson, C. Yu and T. Başar. On the Analysis of the DeGroot-Friedkin Model with Dynamic Relative Interaction Matrices. In *20th IFAC World Congress*, pp. 12408-12413, Toulouse, France, 2017.
- [17] Q. Liu, **M. Ye**, Z. Sun, J. Qin and C. Yu. Coverage Control of Unicycle Agents Under Constant Speed Constraints. In *20th IFAC World Congress*, pp. 2526-2531, Toulouse, France, 2017.

- [18] J. Russell, **M. Ye**, B. D.O. Anderson, H. Hmam and P. Sarunic. Cooperative Localisation of a GPS-Denied UAV in 3-Dimensional Space Using Direction of Arrival Measurements. In *20th IFAC World Congress*, pp. 8291-8296, Toulouse, France, 2017.
- [19] **M. Ye**, B. D.O. Anderson and C. Yu. Model-Independent Trajectory Tracking of Euler-Lagrange Agents on Directed Networks. In *Proc. 55th IEEE Conf. on Decision and Control*, Las Vegas, USA, 2016, pp. 6921-6927.
- [20] Q. Liu, **M. Ye**, J. Qin and C. Yu. Event-Based Leader-Follower Consensus for Multiple Euler-Lagrange Systems with Parametric Uncertainties. In *Proc. 55th IEEE Conf. on Decision and Control*, Las Vegas, USA, 2016, pp. 2240-2246.
- [21] L. Zhang, **M. Ye**, B. D.O. Anderson, P. Sarunic and H. Hmam. Cooperative Localisation of UAVs in a GPS-denied Environment Using Bearing Measurements. In *Proc. 55th IEEE Conf. on Decision and Control*, Las Vegas, USA, 2016, pp. 4320-4326.
- [22] **M. Ye**, B. D.O. Anderson and C. Yu. A Variable Gain Model-Independent Algorithm for Rendezvous of Euler-Lagrange Agents on Directed Networks. In *Australian Control Conference*, Newcastle, Australia, 2016, pp.153-158.
- [23] **M. Ye**, C. Yu and B. D.O. Anderson. Model-Independent Rendezvous of Euler-Lagrange Agents on Directed Networks. In *Proc. 54th IEEE Conf. on Decision and Control*, Osaka, Japan, 2015, pp.3499-3505.

Extended Conference Abstracts

- [24] **M. Ye**, J. Liu, and B. D.O. Anderson. Opinion Dynamics with State-Dependent Susceptibility to Influence. In *23rd Mathematical Theory of Networks and Systems*, pg. 820-823, Hong Kong, July 2018.
- [25] S. Kawamura, K. Cai, and **M. Ye**. Further Study on Parameter Bound of Surplus-Based Averaging Algorithm. In *23rd Mathematical Theory of Networks and Systems*, pg. 787-790, Hong Kong, July 2018

Acknowledgments

This PhD has been a wonderful and defining journey which has been enriched immeasurably by the many people whom I have been fortunate to share it with. First, I want to express my heartfelt gratitude to my supervisors, Brian D.O. Anderson and Brad Yu. They captured my imagination, introducing me to the wonders of research as an undergraduate summer scholar, and it only seems right that I have reach this milestone with their continued guidance. I consider it an absolute privilege to have Brian and Brad as my supervisors.

Beyond the extensive technical training I have received, Brian has provided me with steady guidance on how to develop in all areas as a researcher, including the softer skills of leadership, communication, and management. No PhD is without major difficulties, and the encouragement and support I received from Brian during these points in time has been invaluable. He has always been there to provide sound advice and balance, even beyond matters of research. Through all these things, I have come to comprehend the deep meaning of *Doktorvater*. I must also thank Dianne Anderson, who hosted me several times with deliciously cooked food, and was ever understanding when I arrived at their house to talk more about research!

Brad has been untiring in supporting me as I explore the many wonderful topics within systems and control. The enormous freedom given to me so that I can pursue topics of my choice, including a switch half way through my PhD from conventional multi-agent systems to social network analysis, is deeply appreciated. I am forever grateful for his unwavering confidence and encouragement to allow me such a unique opportunity. More than that, Brad has taught me the importance of having a clear vision about my research, both for the near term problems and how they will build towards a framework that can have meaningful impact in our community and beyond.

There are a number of people in CECS whom I owe many thanks to. Prof. Jochen Trunpf, thank you for teaching me the many nuanced and unwritten aspects of research, teaching, and academia. Each time I talk to you, it seems like a new trick has been added to my toolbox. To Dr. Guodong Shi, thank you for the reflective and thoughtful conversations we had about life as a researcher. Your energy and enthusiasm is boundless, and your imparted wisdom and advice is deeply appreciated. To the lunch time crew (which seems to shrink every passing year) of Dr. Yifei Huang, Alex Martin, YonHon Ng, and Prof. Rajeev Gore: thank you for providing me with a daily break from the research. Aside from our lunchtime exploits in the FOCUS Word Puzzle, I enjoyed our conversations about tennis, Formula One, the housing market, stock investments and anything that did not involve differential or difference equations. In particular, I want to thank YonHon. You have been a most wonderful

friend over the past 4 years. I know that I shared many of my difficulties with you, be they research or otherwise, and you have been always there to listen and give advice. I hope that I have been able to do the same for you.

To Dr. Zhiyong Sun and Dr. Qingchen Liu, thanks for making the lab a fun place to be in! I will miss coming to your office every 30 minutes when I'm bored. I always learned something new (usually research related) talking to Zhiyong, and he has been a constant source of inspiration for me. Qingchen, thanks for your help with everything. I thoroughly enjoyed the papers we co-authored, and the sweat we shed together working towards paper deadlines. To Zhixun Li, Pengfei Fang, Dr. Junming Wei and Dr. Yun Hou, thanks for fostering the group's unity throughout the years; this group means much more to me than just our exploits in research. I will miss our regular watermelon feasts in Zhiyong and Qingchen's office, and the hotpot dinners. I would like to acknowledge the support from the Commonwealth Government, via the Australian Government Research Training Program Scholarship, and thank Prof. Saman Halmagugae for the generous support in the final months of my PhD.

I would also like to thank my co-authors Dr. Minh Hoang Trinh, Dr. Young-Hun Lim, Alain Govaert, and Yuzhen Qin. It has been a pleasure tackling interesting and challenging problems with you. In particular, you have made me recognise just how much can be achieved through collaboration, and I know that I have only reached this far thanks to your invaluable help. Thanks also for being such great hosts when I visited your lab groups, which gave me countless memories to cherish. To Prof. Ji Liu, thanks for the close guidance and patience, and the many enthusiastic discussions we shared on the topic of opinion dynamics. To Dr. Hatem Hmam, Dr. Sam Drake, and Dr. Peter Sarunic of the Australian Defence Science and Technology Group, thank you for the frequent interactions we had on relevant and interesting problems. Being able to work on applications based topics for Defence provided excellent variation to the research on social networks.

One of the most wonderful aspects of research is the number amazing people I have had the fortune of meeting and interacting with, both at conferences and while visiting various labs. In particular, I would like to mention Dr. Na Huang, Dr. Chao Huang, Yaoxian Song, Li Gao, Dr. Tingrui Han, Dr. Xiaocheng Wang, Dr. Peng Wang, Chuong Van Nyugen, Phuong Hoang, Pham Hoang Viet, Carlo Cendese, Dr. Wenjun Mei, Dr. Qingkai Yang, Dr. Yuke Li and many others. You've all been incredibly helpful during my travels and visits, making each trip an absolute joy. I am truly grateful for your friendship.

To Prof. Zhiyun Lin (Hangzhou Dianzi Uni, China), Prof. Wei Ren (UC Riverside, USA), Prof. Kai Cai (Osaka City Uni, Japan), Prof. Ming Cao (Uni of Groningen, Netherlands), and Prof. Hyo-Sung Ahn (GIST, Korea), thank you for generously hosting me at your labs. I am incredibly privileged and grateful for the opportunities to immerse myself in your research groups and develop my knowledge, learning new and valuable tools to approach research problems from different perspectives. Your wisdom and guidance will remain with me forever. To other members of the community whom I have had the pleasure of interacting with at conferences, Prof. Tamer Başar, Prof. Julien Hendrickx, Prof. Steve Morse, Prof. Shaoshuai Mou,

Prof. Francesco Bullo, Prof. Jiahu Qin, and many more, thank you for providing me with exciting new ideas and insightful discussions about a great many topics. It has helped fuel my enthusiasm for systems and control research and you all continue to inspire me to keep chasing my goals.

To the badminton community of Canberra, thanks for giving me a place to escape and refresh and grow my love for the best sport in the world. In particular, I want to thank Rita Chou, Johann Kwan, Victoria Wang, Kevin Ko, Kola Khamchaleun, Jay Cruise, Xing Huong Goh, Xing Yang Goh, Jules Smith, Eva Wang, Josiah Li, Tang, the Souksovat family (Pascal, Cedric, Aline, Val). You have all gifted me with a tremendous amount of joy and helped me stay passionate about badminton. To Nathan, thanks for all the laughs, and to Lawrence, thanks for helping me develop my patience. To Melanie Pill, thanks for your kindness and advice, and to Seng Low, thanks for the great catch phrases and free socks. To Bong Wong, and Sheehan Lim, I've enjoyed every moment of our friendship and all the stupid things we have done. I will cherish our training adventures in Taiwan for years to come. To Terrence Wong, Edward Wang, Daniel Lee, Tomohiro Kinoshita, and Anthony Joe, thanks for the wonderful memories of the 2015 Australian University Games; a gold medal is an unexpected but most welcome surprise to my PhD journey.

To the Joe family: thank you for welcoming me with open arms. Canberra has felt like my home over the past 4 years in no small part because of your kindness and generosity. To Anthony, thanks showing me that it's possible to become an international athlete without ever really having to grow up. I've enjoyed, and always take inspiration from, your badminton exploits. To Shuen and Kam, thanks for looking after me and taking great care of me. A delicious dinner certainly helped me reset on those days where progress was slow at the lab. To Jacinta, thank you for being with me every step of the way. I am forever grateful for your boundless patience and understanding and love, and the unfailing ability to make me laugh when I most needed it. Over these four years, you have been the rock that I always turned to in order to draw strength, and find stability.

Last, I would like to thank my family, both in China and in New Zealand. I cannot overstate how much it means to me that you are always there with open arms when I come back home (however infrequently that may be). I hope that I will continue to make you all proud. In particular, I want to thank my parents, Yan Xiong and Xiao Qing, who have been incredibly supportive of me throughout this endeavour. With each step in my career, I seem to be moving further from our wonderful home in New Zealand. You have allowed me to chase my dreams with freedom, providing me with unconditional love and encouragement. Without your love and care over the many years, I would not be the person I am today, and this thesis would not have been possible.

Abstract

A fundamental aspect of society is the exchange and discussion of opinions between individuals, occurring in mediums and situations as varied as company boardrooms, elementary school classrooms and online social media. This thesis studies several mathematical models of how an individual's opinion(s) evolves via interaction with others in a social network, developed to reflect and capture different socio-psychological processes that occur during the interactions.

In the first part, and inspired by Solomon E. Asch's seminal experiments on conformity, a novel discrete-time model of opinion dynamics is proposed, with each individual having both an expressed and a private opinion on the same topic. Crucially, an individual's expressed opinion is altered from the individual's private opinion due to pressures to conform to the majority opinion of the social network. Exponential convergence of the opinion dynamical system to a unique configuration is established for general networks. Several conclusions are established, including how differences between an individual's expressed and private opinions arise, and how to estimate disagreement among the private opinions at equilibrium. Asch's experiments are revisited and re-examined, and then it is shown that a few extremists can create "pluralistic ignorance", where *people believe* there is majority support for a position but in fact the position is *privately rejected* by the majority of individuals!

The second part builds on the recently proposed discrete-time DeGroot–Friedkin model, which describes the evolution of an individual's self-confidence (termed *social power*) in his/her opinion over the discussion of a sequence of issues. Using nonlinear contraction analysis, exponential convergence to a unique equilibrium is established for networks with constant topology. Networks with issue-varying topology (which remain constant for any given issue) are then studied; exponential convergence to a unique limiting trajectory is established. In a social context, this means that each individual forgets his/her initial social power exponentially fast; in the limit, his/her social power for a given issue depends only on the previously occurring sequence of dynamic topology. Two further related works are considered; a network modification problem, and a different convergence proof based on Lefschetz Fixed Point Theory.

In the final part, a continuous-time model is proposed to capture simultaneous discussion of *logically interdependent topics*; the interdependence is captured by a "logic matrix". When no individual remains attached to his/her initial opinion, a necessary and sufficient condition for the network to reach a consensus of opinions is provided. This condition depends on the interplay between the network interactions and the logic matrix; if the network interactions are too strong when compared to the logical couplings, instability can result. Last, when some individuals remain attached to their initial opinions, sufficient conditions are given for opinions to converge to a state of persistent disagreement.

Contents

Declaration	vii
Acknowledgments	xi
Abstract	xv
1 Introduction	1
1.1 Social Network Analysis	1
1.2 Opinion Dynamics and Influence Networks	2
1.2.1 The Fundamental French–Harary–DeGroot Model	2
1.2.2 Beyond Consensus and Towards Complex Social Phenomena	3
1.2.3 Social Phenomena of Relevance	5
1.2.4 Motivations and Key Concepts	6
1.2.5 Relation to Coordination of Multi-Agent Systems	7
1.3 Thesis Outline and Statements of Collaborations	8
2 Preliminaries	11
2.1 Notations and Definitions	11
2.2 Graph Theory	12
2.3 Basic Models of Opinion Dynamics	13
2.3.1 The DeGroot Model	13
2.3.2 The Friedkin–Johnsen Model	16
2.3.3 Comments on the Models	17
I How Differences in Private and Expressed Opinions Arise	19
3 A Novel Model for Opinion Dynamics Under Pressure to Conform	21
3.1 Introduction	22
3.1.1 Chapter Organization	23
3.2 A Model with Expressed and Private Opinions	23
3.2.1 The Opinion Dynamics Model	23
3.2.2 Local Public Opinion	27
3.2.3 Obtaining a Compact Form for the Influence Network	28
3.3 Convergence Properties of the Model	30
3.3.1 Causes of Persistent Disagreement and Differences in Opinions	32
3.3.2 Estimating Disagreement in the Private Opinions	35

3.3.3	An Individual's Resilience Affects Everyone	36
3.3.4	Local Public Opinions	39
3.3.5	Simulations	40
3.4	Conclusion	42
3.5	Appendix: Proofs and Simulations	42
3.5.1	Proof of Lemma 3.1	42
3.5.2	Proof of Lemma 3.2	43
3.5.3	Proof of Lemma 3.3	44
3.5.4	Simulation Counter-Example	45
3.5.5	Proof of Corollary 3.2	46
3.5.6	Simulation Influence Matrix	48
4	The EPO Model's Connections with Social Psychology Concepts	51
4.1	Introduction	51
4.1.1	Introduction to Asch's Experiments	51
4.1.2	Pluralistic Ignorance in Social Networks	52
4.1.3	Chapter Organization	54
4.2	Investigation of Asch's Conformity Experiments	54
4.2.1	Theoretical Analysis	56
4.2.2	Simulations	58
4.3	A Few Zealots Can Create Pluralistic Ignorance	62
4.3.1	Base Simulation Set-Up	63
4.3.2	Simulation Set-Up and Results For Small-World Networks	65
4.3.3	Simulation Set-Up and Results For Scale-Free Networks	66
4.3.4	Discussion for Small-World Networks	71
4.3.5	Discussion for Scale-Free Networks	75
4.3.6	Key Observations and Insights	76
4.4	Conclusions	78
<hr/>		
II	Evolution of Individual Social Power	79
5	Evolution of Social Power in Networks with Constant Topology	81
5.1	Introduction	81
5.1.1	The DeGroot–Friedkin Model	83
5.1.1.1	DeGroot Consensus of Opinions	84
5.1.1.2	Friedkin's Self-Appraisal Model for Determining Self-Weight	84
5.1.1.3	Existing Results	86
5.1.2	Chapter Organization	88
5.2	Exponential Convergence to Constant Social Power	89
5.3	Further Analysis of Dynamical Behaviour	95
5.3.1	A Contraction-Like Property	95
5.3.2	Upper Bound on Individual's Social Power at Equilibrium	96

5.3.3	Convergence Rates	97
5.4	Conclusions	98
6	Dynamic Social Networks: Exponential Forgetting of Perceived Social Power	99
6.1	Introduction	99
6.1.1	Motivating Examples for Issue-Varying Topology	101
6.1.2	Chapter Organization	102
6.2	The Dynamic Topology Model and Objective	102
6.3	Exponential Convergence to a Unique Limiting Trajectory	104
6.3.1	Extending the Results in Section 5.3	107
6.4	Periodically Varying Topology	108
6.5	Simulations	112
6.6	Conclusions	113
6.7	Appendix: Simulation Parameters	116
7	Modification of Social Dominance in Autocratic Networks	117
7.1	Introduction	117
7.1.1	Chapter Organization	119
7.2	Problem Set-up and Proposed Modification Strategies	119
7.3	Results and Their Social Interpretations	123
7.3.1	Topology Variation 7.1: Single Attack	125
7.3.2	Topology Variation 7.2: Coordinated Double Attack	126
7.3.3	Topology Variation 7.3: Uncoordinated Double Attack	127
7.3.4	Topology Variation 7.4: Two Dissenting Subjects	128
7.3.5	Topology Variation 7.5: Leadership Group	129
7.4	Simulations	130
7.5	Conclusions	131
7.6	Appendix: Proofs	133
7.6.1	Proofs for Theorem 7.2 and Corollary 7.1	133
7.6.2	Proofs for Theorem 7.3 and Corollary 7.2	133
7.6.3	Proof for Theorem 7.4	134
7.6.4	Proof for Theorem 7.5	135
7.6.5	Proof for Theorem 7.6	136
8	Nonlinear Mapping Convergence and Application to Social Power Analysis	137
8.1	Introduction	137
8.1.1	Chapter Organization	139
8.2	A General Nonlinear Mapping Convergence Result	139
8.2.1	Background on Lefschetz Fixed Point Theory	139
8.2.2	A General Convergence Result	141
8.3	Application to the DeGroot-Friedkin Model	142
8.3.1	Existing Results	143

8.3.2	Proof of a Unique Fixed Point Which Is Locally Exponentially Stable	144
8.4	Conclusions	148

III	Opinion Dynamics with Interdependent Topics	149
------------	--	------------

9	Continuous-Time Opinion Dynamics with Interdependent Topics	151
9.1	Introduction	151
9.1.1	Chapter Organization	153
9.2	Development of the Continuous-Time Model	153
9.2.1	Opinion Dynamics Model and Problem Statement	154
9.2.2	Interdependent Topics and the Logic Matrix	156
9.3	Consensus for Individuals with No Stubbornness	157
9.3.1	Stability and Convergence to Consensus	158
9.3.2	Consensus for a Class of Logic Matrices	163
9.4	Persistent Disagreement for Stubborn Individuals	165
9.4.1	Stability for Social Networks Under Assumption 9.1	166
9.4.1.1	Individuals Are Slightly Stubborn	166
9.4.1.2	All Individuals Are Extremely Stubborn	167
9.4.1.3	Individuals Have Almost the Same Stubbornness	167
9.4.2	Stability for Social Networks Under Assumption 9.2	168
9.5	Simulations	169
9.6	Conclusions	170
9.7	Appendix: Some Proofs	173
9.7.1	Proof of Lemma 9.1	173
9.7.2	Relation to Discrete-Time Model	174

10	Conclusions and Future Work	177
10.1	Thesis Summary and Contributions	177
10.1.1	How Differences in Private and Expressed Opinions Arise	177
10.1.2	Evolution of Individual Social Power	178
10.1.3	Opinion Dynamics with Interdependent Topics	179
10.2	Future Work	180
10.2.1	The EPO Model	180
10.2.2	The DeGroot–Friedkin Model	181
10.2.3	Multiple Logically Interdependent Topics	182
10.2.4	Longer-Term Research Problems: Investigation of New Social Phenomena	183

A	Appendix	185
A.1	Linear Algebra	185
A.1.1	Consensus Level and Coefficient of Ergodicity	187
A.2	Graph Theory	187

A.3 Nonlinear Contraction Analysis	189
--	-----

List of Figures

1.1	The structure of the thesis chapters and the models studied.	9
2.1	An example social network of 5 individuals modelled as the directed graph $\mathcal{G}[\mathbf{W}]$. The direction of the edges can be considered to represent influence flow. For example, individual 1 exerts a direct influence onto individual 2 to change 2's opinion via the directed edge e_{12} , with influence weight w_{21} . Individual 1 indirectly influences individual 3 via the directed path formed by the sequence of edges e_{15}, e_{54}, e_{43} , with weights w_{51}, w_{45}, w_{34}	14
3.1	Each individual undergoes this process of influence and opinion evolution, occurring simultaneously with all other individuals. At time step t , individual i expresses opinion $\hat{y}_i(t)$ and learns of i 's neighbours' expressed opinions $\hat{y}_j(t), j \in \mathcal{N}_i$. This might occur in a boardroom meeting or in an online social medium. Next, individual i 's privately held opinion evolves to be $y_i(t+1)$ under the influence of i 's own privately held opinion $y_i(t)$ and the expressed opinions of i 's neighbours, $\hat{y}_j(t)$. The precise algorithm is given in Eq. (3.1). Once individual i has updated his/her privately held opinion, i then determines the new $\hat{y}_i(t+1)$ to be expressed in the next round of discussion, as given in Eq. (3.2).	24
3.2	Individual i 's private opinion $y_i(t)$ and expressed opinion $\hat{y}_i(t)$ are represented by the blue node and purple node, respectively. The arrows/edges represent quantities which exert an influence on i 's private or expressed opinion. In particular, $y_i(t)$ is changed due to interpersonal influence from other individuals in the influence network, i 's self-weight (if there is a self-loop), and an attachment to initial opinion $y_i(0)$. The private opinion of individual i exerts an influence on i 's expressed opinion via a resilience to conformity, while the public opinion \hat{y}_{avg} exerts a pressure to conform. Individual i exerts an influence on others in the network via i 's expressed opinion \hat{y}_i	25

3.3	The influence network $\mathcal{G}[\mathbf{W}]$ of n individuals, with each individual's private and expressed opinions evolving according to Eq. (3.1) and Eq. (3.2), is equivalent to the influence network $\mathcal{G}[\mathbf{P}]$ with $2n$ nodes, evolving according to the system dynamics Eq. (3.5). Nodes $\mathcal{V}_p = \{v_1, \dots, v_n\}$ are associated with the private opinions \mathbf{y} , and induce the <i>private opinion subgraph</i> $\mathcal{G}[\mathbf{P}_{11}]$. Nodes $\mathcal{V}_e = \{v_{n+1}, \dots, v_{2n}\}$ are associated with the expressed opinions $\hat{\mathbf{y}}$, and induce the <i>expressed opinion subgraph</i> $\mathcal{G}[\mathbf{P}_{22}]$. The off-diagonal terms \mathbf{P}_{12} and \mathbf{P}_{21} describe influences (edges) between nodes in \mathcal{V}_p and \mathcal{V}_e . Here, an illustrative example with $n = 3$ is shown.	29
3.4	Temporal evolution of opinions for 18 individuals in an influence network. The green and dotted blue lines represent the expressed and private opinions of the individuals, respectively.	41
3.5	Temporal evolution of opinions for 18 individuals in an influence network. The green and dotted blue lines represent the expressed and private opinions of the individuals, respectively. The lack of stubbornness, $\lambda_i = 1, \forall i$, means that all opinions reach a consensus.	41
3.6	Evolution of opinions for counter-example to nonexpansive dynamics.	46
4.1	Example of the Asch experiment, with an illustrative influence network of eight individuals in the Asch. Here, the edges are bidirectional (for convenience only), indicating both individuals influence one another. The individuals are required to openly discuss the length of lines, and state their individual beliefs as to which one of A, B, C has the same length as the green line. Clearly A is equal in length to the green line. The confederates (seven blue individuals) unanimously express belief in the same wrong answer, e.g. B	53
4.2	The function $f(\lambda_1)$ and $1 - \lambda_1$ plotted against λ_1 . The analytical calculations show that $y_1^* = f(\lambda_1)$, and thus the red line represents individual 1's final private belief as a function of his susceptibility to influence.	59
4.3	The function $g(\phi_1, n)$, with $n = 2, 4, 8$, plotted against ϕ_1 . The analytical calculations show that $\hat{y}_1^* = g(\phi_1, n)y_1^*$, and thus the plot shows how the test individual's final expressed opinion is changed from his final private opinion by his resilience ϕ_1 , and by n	59
4.4	Behaviour of the test individual as a function of susceptibility to influence, λ_1 , and resilience to pressure, ϕ_1	60
4.5	Evolution of beliefs for an <i>independent</i> individual, $\lambda_1 = 0.1, \phi_1 = 0.9$	60
4.6	Evolution of beliefs for a <i>yielding</i> individual with <i>distortion of judgment/perception</i> , $\lambda_1 = 0.9, \phi_1 = 0.1$	61
4.7	Evolution of beliefs for a <i>yielding</i> individual with <i>distortion of action</i> , $\lambda_1 = 0.1, \phi_1 = 0.1$	61
4.8	Histogram example of distribution of civilian resilience, ϕ_i following a beta distribution with $\alpha = 2, \beta = 15$	65

4.9	Mean of the degree distributions for the small-world networks in the simulations.	66
4.10	Evolution of private opinions, $y_i(t)$, over time t , for an example small-world network with $n = 200$ and 20 zealots. The red bold line at $y = 1$ represents the zealots' opinions. The other coloured lines represent civilian opinions.	67
4.11	Evolution of expressed opinions, $\hat{y}_i(t)$, over time t , for an example small-world network with $n = 200$ and 20 zealots. The red bold line at $y = 1$ represents the zealots' opinions. The other coloured lines represent civilian opinions.	67
4.12	Distribution of civilians' final private and expressed opinions, y_c^* and \hat{y}_c^* , with 0 zealots in $n = 200$ small-world network.	68
4.13	Distribution of civilians' final private and expressed opinions, y_c^* and \hat{y}_c^* , with 1 zealot in $n = 200$ small-world network.	68
4.14	Distribution of civilians' final private and expressed opinions, y_c^* and \hat{y}_c^* , with 5 zealots in $n = 200$ small-world network.	69
4.15	Distribution of civilians' final private and expressed opinions, y_c^* and \hat{y}_c^* , with 10 zealots in $n = 200$ small-world network.	69
4.16	Distribution of civilians' final private and expressed opinions, y_c^* and \hat{y}_c^* , with 20 zealots in $n = 200$ small-world network.	70
4.17	Example degree distribution of a scale-free network used in the simulations.	71
4.18	Evolution of private opinions, $y_i(t)$, over time t , for an example scale-free network with $n = 200$ and 5 zealots, using targeted zealot placement , and local public opinion . Notice the similarity to Fig. 4.10 because civilians typically have low susceptibility λ_i . The red bold line at $y = 1$ represents the zealots' opinions. The other coloured lines represent civilian opinions.	72
4.19	Evolution of expressed opinions, $\hat{y}_i(t)$, over time t , for an example scale-free network with $n = 200$ and 5 zealots, using targeted zealot placement , and local public opinion . The red bold line at $y = 1$ represents the zealots' opinions. The other coloured lines represent civilian opinions.	72
4.20	Evolution of expressed opinions, $\hat{y}_i(t)$, over time t , for an example scale-free network with $n = 200$ and 5 zealots using targeted zealot placement , and global public opinion . The red bold line at $y = 1$ represents the zealots' opinions. The other coloured lines represent civilian opinions.	73
4.21	Distribution of civilians' final private and expressed opinions, y_c^* and \hat{y}_c^* , with 5 zealots in $n = 200$ scale-free network, using targeted zealot placement scheme and local public opinion	73
4.22	Distribution of civilians' final private and expressed opinions, y_c^* and \hat{y}_c^* , with 5 zealots in $n = 200$ scale-free network, using targeted zealot placement scheme and global public opinion	74

4.23	Distribution of civilians' final private and expressed opinions, y_c^* and \hat{y}_c^* , with 5 zealots in $n = 200$ scale-free network, using random zealot placement scheme and local public opinion	74
4.24	The key ways a zealot is able to influence the civilian population for different types of networks are illustrated here. Note that while the edges in the simulations are bidirectional, for clarity, only some key unidirectional edges are shown here.	77
6.1	Evolution of individuals' social powers $x(s)$ for initial condition vector $\hat{x}(0)$. Viewed in conjunction with Fig. 6.2, it is clear that each individual i 's social power trajectory $x_i(s)$ converges to $x_i^*(s)$ by about $s = 10$	113
6.2	Evolution of individuals' social powers $x(s)$ for initial condition vector $\tilde{x}(0)$. Viewed in conjunction with Fig. 6.1, it is clear that each individual i 's social power trajectory $x_i(s)$ converges to $x_i^*(s)$ by about $s = 10$	114
6.3	Evolution of selected individuals' social powers $x_i(s)$, for $i = 1, 3, 6$. The trajectory $x_i(s)$, for $i = 1, 3, 6$, beginning from the two different initial conditions, converges to the same trajectory by about $s = 10$, which is precisely the unique limiting trajectory $x_i^*(s)$	114
6.4	Evolution of individuals' social powers $x(s)$ for initial condition vector $\hat{x}(0)$, with periodically-varying topology. Clearly, $x(s)$ becomes a periodic sequence by about $s = 10$	115
6.5	Evolution of selected individuals' social powers $x_i(s)$: a comparison of different initial condition vectors $\hat{x}(0)$ and $\tilde{x}(0)$, with periodically-varying topology. The trajectory $x_i(s)$, for $i = 1, 3, 6$, beginning from the two different initial conditions, converges to the same periodic sequence $x_i^*(s)$ by about $s = 10$	115
7.1	Star Topology with red centre node v_1 , with $n = 7$ total individuals.	119
7.2	Topology Variation 7.1 (Single Attacker) with $n = 8$, attacker nodes are green.	124
7.3	Topology Variation 7.2 (Coordinated Double Attacker) with $n = 9$, attacker nodes are green.	124
7.4	Topology Variation 7.3 (Uncoordinated Double Attacker) with $n = 9$	124
7.5	Topology Variation 7.4 (Two Dissenting Subjects) with $n = 7$	125
7.6	Topology Variation 7.5 (Leadership Group) with $n = m = 4$	125
7.7	Simulation of Topology Variation 7.1 with $n = 8$, showing select individual $i = 1, 2, 7, 8$'s social powers at equilibrium, x_i^* as a function of the weight β , i.e. the relative interpersonal weight accorded to v_8 by v_1 . As β increases, x_1^* decreases. When $\beta > 1 - c_{17}$, v_1 is no longer the socially dominant individual in the network, as predicted by Theorem 7.2.	131

7.8	Simulation of Topology Variation 7.4 with $n = 8$, showing select individual $i = 1, 2, 7, 8$'s social powers at equilibrium, x_i^* as a function of the weight β_1 , i.e. the relative interpersonal weight accorded to v_8 by v_7 . In this instance, $\beta_2 < \sum_{i=2}^{n-2} c_{1,i}$, which does not satisfy the inequality on β_2 in Statement (ii), Theorem 7.5. No matter how β_1 varies, $x_1^* > x_i^*$ for all $i \neq 1$	132
7.9	Simulation of Topology Variation 7.4 with $n = 8$, showing select individual $i = 1, 2, 7, 8$'s social powers at equilibrium, x_i^* as a function of the weight β_1 , i.e. the relative interpersonal weight accorded to v_8 by v_7 . In this instance, $\beta_2 > \sum_{i=2}^{n-2} c_{1,i}$, i.e. the inequality on β_2 in Statement (ii), Theorem 7.5 is satisfied. Thus, when β_1 is sufficiently large (see Statement (ii), Theorem 7.5), $x_1^* < x_8^*$ and centre v_1 is no longer the socially dominant individual.	132
9.1	Consensus is reached on all 3 topics when the conditions of Theorem 9.1 are satisfied.	171
9.2	When the topics are uncoupled, $C = I_d$, consensus is reached but the final consensus values are different due to the lack of logic coupling. . .	171
9.3	When Condition 9.11 of Theorem 9.1 is not satisfied, the opinion system is unstable.	172
9.4	In the presence of four stubborn individuals, a state of persistent disagreement is achieved when the conditions in Theorem 9.4 are met. . .	172
A.1	Two trajectories of $\mathbf{x}(k+1) = \mathbf{f}_{\sigma(k)}(\mathbf{x}(k), k)$ are separated by a virtual displacement $\delta\mathbf{x}(k)$. The virtual dynamics are illustrated, and nonlinear contraction analysis is focused on study of how the virtual displacement evolves over k . One desirable property is for the virtual displacement $\delta\mathbf{x}(k)$ to converge to zero; this implies certain properties of the system A.13. See Chapters 5 and 6 for the interpretations of these properties in the context of the evolution of social power in social networks.	190

Introduction

1.1 Social Network Analysis

The works of this thesis fall into the broad field of research on social networks which has been studied, in different communities at various times, for many decades. A general definition of a social network is a system which contains a set of social actors (the actor may be an individual or an organisation) that interact according to a set of social relationships or interconnections. Social network analysis is the study of variables of interest in the network, e.g. the actors' opinions on a given topic, and how these variables may be determined or may be changed due to the interactions. In many, but not all of the works, a focus is placed on investigating the role of network itself in shaping the behaviour that arises.

From one perspective, social network analysis falls into the rich and vast field of complex network analysis; by nature, humans react in complex and often surprising ways to stimulus when in a social setting. The spreading of rumours or ideas over a social network can be captured by a diffusion model [Kempe et al., 2003], or by epidemic models which first arose in the study of virus spreading [Mei et al., 2017; Brockmann and Helbing, 2013]. Investigations have been made regarding the differences in how Facebook users consume scientific and conspiracy articles [Del Vicario et al., 2016]. The work of [Ramazi et al., 2016] studied networks of binary decision making individuals, classifying individuals as either conforming (the individual takes decision A when enough neighbours are also taking A) or anti-conforming (the individual takes decision B when too many neighbours take A). Surprisingly, it was shown that networks of either conforming or anti-conforming individuals tend to reach decisions they are satisfied with. Systematic changing of the network structure to undermine or strengthen the network with respect to an objective function, e.g. reducing the expected number of attacks arising from a terrorist network, was studied in [Mellon et al., 2016]. A comprehensive review of literature on social network analysis is well beyond the scope of this thesis; the above is just a selective representation that helps to highlight the wide range of interesting problems considered.

One of the most popular approaches to social network analysis is agent-based modelling. In an agent-based model, each actor is represented by an agent, with some mathematical model (typically a vector difference or differential equation) describing the agent's dynamics, e.g. how an individual may process new sources of

information but with a confirmation bias towards the individual's established views. Agents interact according to mathematical models which describe rules or processes that occur during the interaction. Beyond building a model, a key task is to establish what types of network-level phenomena (macro-level dynamics) arise from processes at the agent level (micro-level dynamics); some resulting network-level phenomena can be counter to what is expected when considering what occurs at the agent-level dynamics. This thesis will study agent-based models of opinion dynamics.

1.2 Opinion Dynamics and Influence Networks

Opinion dynamics is an area of research that has been investigated by communities in sociology, physics, computer science, systems and control engineering, and more, over decades. Thus, no attempt is made to give a detailed account of all available works other than to introduce some of the most relevant results. At the beginning of each main chapter, a more detailed review may be provided to give relevance and context to each work. The interested reader is referred to [Proskurnikov and Tempo, 2017] for an overview survey from a systems and control perspective, [Flache et al., 2017] for a survey from the computational sociology community while the survey [Friedkin, 2015] forms a bridge between the two communities. In the remainder of this section, the fundamental French–Harary–DeGroot model is highlighted and some interesting extensions, which aim to capture a variety of social phenomena, are discussed. Models and social phenomena of direct interest to this thesis are explored, then motivations and key philosophies of opinion dynamics research are elucidated, and last, a connection is drawn to multi-agent systems research.

1.2.1 The Fundamental French–Harary–DeGroot Model

In 1956, John French Jr. introduced an agent-based model of opinion dynamics [French Jr, 1956] to study how individuals exerted social power on each other during interactions in a network. French Jr. assumed that each individual has an opinion on a given topic, and each individual interacts to learn of the opinions of that individual's neighbours. In doing so, the opinions evolve over time as each individual integrates learned opinion values of his/her neighbours with the individual's own opinion until a consensus is reached on the opinion value, i.e. there is agreement across the opinions of all individuals. Frank Harary [Harary, 1959; Harary et al., 1965] and Morris DeGroot [DeGroot, 1974] provided major theoretical contributions to generalise the model and establish conditions on the network which guarantee a consensus of opinions is achieved. The French–Harary–DeGroot model has become the fundamental agent-based model of opinion dynamics which many subsequent works, including the works in this thesis, build upon. For convenience, the author will refer to the model simply as the DeGroot model (as it is widely known), but the contributions by French and Harary are not overlooked. To better understand the literature review and place context on results, the model will be informally introduced here, with a formal account given in Chapter 2.

For a given population of n individuals, individual i has an opinion $y_i(t)$ at time $t = 0, 1, \dots$, which is a real-valued number¹. Individual i 's opinion evolves according to the difference equation

$$y_i(t+1) = \sum_{j=1}^n w_{ij} y_j(t) \quad (1.1)$$

and this process occurs simultaneously for all n individuals in the network. Here, $y_j(t)$ is the opinion value of individual j , who is a neighbour of individual i . The nonnegative coefficient w_{ij} is assumed to satisfy $\sum_{j=1}^n w_{ij} = 1$. One can consider the quantity w_{ij} as representing the amount of relative influence individual j exerts on individual i in determining i 's new opinion value $y_i(t+1)$. Thus, w_{ij} is termed an “influence weight” while the collection of individuals is sometimes termed an “influence network”. It is clear that the influence weights represent a social power exerted by one individual onto another (French had been driven to develop a mathematical representation of social power). As proved in [Harary, 1959; Harary et al., 1965; DeGroot, 1974], the opinions asymptotically reach a consensus if the graph representing the network is strongly connected² and aperiodic. That is, $\lim_{t \rightarrow \infty} y_i(t) = y_j(t)$ for all individuals i and j in the network; interpersonal social influence acts to bring opinions closer together in value. Notice that because it was assumed that $\sum_{j=1}^n w_{ij} = 1$, then $y_i(t+1)$ is a convex combination (or weighted average) of opinions $y_j(t)$. Thus, some works refer to such models as weighted averaging models.

The DeGroot model has received extensive treatment since the seminal works discussed above. A number of experimental validations have been conducted, e.g. [Becker et al., 2017; Chandrasekhar et al., 2012]. Attention has also been placed on establishing conditions on the network topology which guarantees a consensus is reached even when topology is changing over time, i.e. the influence weights $w_{ij}(t)$ are time-varying. Some results are given in e.g. [Cao et al., 2008; Ren and Beard, 2005; Nedić and Liu, 2017; Shi and Johansson, 2013]. A major assumption of the DeGroot model is that all individuals *simultaneously* update their opinions at each time instant t . Gossip-based models relax this assumption. Gossip models, in which only one individual's opinion changes at each time instant (or a subset of individuals in some variations), have been studied in e.g. [Ravazzi et al., 2015; Liu et al., 2011; Boyd et al., 2006; Yu et al., 2017]. A continuous-time variant was proposed in [Abelson, 1964].

1.2.2 Beyond Consensus and Towards Complex Social Phenomena

Beyond consensus, variations of the DeGroot model have been proposed to investigate how different social phenomena may arise. The Hegselmann–Krause model [Hegselmann and Krause, 2002; Lorenz, 2007; Blondel et al., 2009; Mirtabatabaei and Bullo, 2012; Su et al., 2017; Etesami and Başar, 2015] introduced the concept of bounded-confidence to describe homophily, where an individual interacts only

¹Other works may assume the opinion is a discrete variable. An interpretation of an opinion as a real number is provided in Chapter 2.

²This condition implies that for any two individuals i and j , i can directly or indirectly (via a path on the graph) influence j 's opinion. The reader is referred to Chapter 2 for details.

with those others who have similar opinions. From a modelling perspective, this means that the influence weight w_{ij} is dependent on the difference $|y_i(t) - y_j(t)|$, i.e. w_{ij} is state-dependent. In the limit, individuals can become separated into clusters, where the final opinions are the same within each cluster, and different between the clusters. The Altafini model introduced the concept of negative influence weights $w_{ij} < 0$ to capture antagonistic interactions among individuals who may, for any number of reasons, dislike or mistrust each other [Altafini, 2013; Proskurnikov et al., 2016; Liu et al., 2017a]. If the strongly connected network is “structurally balanced”, the opinions can become polarised into two opposing clusters, otherwise all opinions converge to $y_i(t) = 0$. The paper [Dandekar et al., 2013] extended the DeGroot model to capture an individual’s tendency for bias assimilation, i.e. an individual will place more weight on opinions whose values are closer to the individual’s current opinion. Under certain conditions on the network structure and intensity of bias assimilation, the social network can become polarised. Social influence, as described by the DeGroot model, was used to explain how the “wisdom of the crowd” effect can be strengthened due to interactions between individuals [Becker et al., 2017], with experimental validation provided. Models which consider an individual’s desire to be unique are studied in [Mäs et al., 2014; Smaldino and Epstein, 2015].

Many of the above listed works contain models (e.g. Hegselmann–Krause and Altafini) which have limiting behaviour that is said to show *weak diversity* [Duggins, 2017; Mäs et al., 2014]. That is, the limiting opinions are diverse (in the sense that the opinions are not at a complete consensus) but form clusters in which every individual in the cluster has the same opinion. In the context of the Hegselmann–Krause model, the network forms disconnected subgraphs where each subgraph represents a cluster, and there is no interaction between individuals in different clusters. In the Altafini model, “structural balance” may be destroyed by changing the sign of a single influence weight from $w_{ij} > 0$ to $w_{ij} < 0$ (or vice versa), with the result being opinions then converge to the neutral value of $y_i(t) = 0$ instead of forming two polarised clusters. In other words, the polarisation phenomenon is not robust to network changes. The reality is that in most social settings, *strong diversity* is observed, where there is a distribution of opinions over a continuous spectrum that is not concentrated into sharp clusters. There may be several subgroups, including perhaps extremists (termed subcultures in the sociology community [Mäs et al., 2014]) where opinions within a subgroup are approximately, but not exactly equal. This has motivated several models which attempt to give different explanations for strong diversity. Works such as [Duggins, 2017; Mäs et al., 2014] show strong diversity can result from a state-dependent intolerance of differing opinions. In their models, $w_{ij}(t)$ is a smooth function which obeys $w_{ij}(t) < 0$ if $|y_i(t) - y_j(t)|$ is greater than some threshold, and $w_{ij}(t) > 0$ if $|y_i(t) - y_j(t)|$ is smaller than that threshold. This contrasts with the Altafini model, which assumes that w_{ij} has a sign that is independent of the state y_i (but may be time-varying).

The Friedkin–Johnsen model [Friedkin and Johnsen, 1990] is a generalisation of the DeGroot model which explains that strong diversity arises due to an individual’s stubbornness, manifested in their remaining to some degree attached to their initial

opinion (dependent on just how stubborn they are). That is, individual i continues to be influenced by $y_i(0)$ over the course of the opinion discussion. From another perspective, this implies that individual i is not maximally susceptible to interpersonal influence as described by the DeGroot model. In the limit, the opinions may converge to an equilibrium of persistent disagreement. A formal introduction to the Friedkin–Johnsen model, which will be used as a basis for some of the results in this thesis, will be given in Chapter 2. The model has been experimentally verified in small- to medium-size groups [Friedkin and Johnsen, 2011; Friedkin and Bullo, 2017], and a continuous-time version of the model appeared earlier [Taylor, 1968]. An interesting extension was presented in [Parsegov et al., 2017], which considers the scenario where individuals simultaneously discuss multiple logically interdependent topics. The final opinion distribution can be heavily affected by the logical interdependence; the paper [Friedkin et al., 2016b] showed that such interdependence could be used to explain why the US public supported the invasion of Iraq in 2003 at the time it occurred, but by 2005 (and as a direct consequence of many individuals changing their opinions concerning the possession or otherwise by Iraq of weapons of mass destruction) the majority of the US public believed the invasion was unjustified.

1.2.3 Social Phenomena of Relevance

The above results give a general overview of results that help to build a picture of the direction this thesis aims to take. Now, a brief introduction is given to the specific social phenomena that are investigated in this thesis; where needed, additional literature is provided in the introduction of the relevant chapter for further in-depth exploration.

In Part I, the causes of discrepancies between an individual’s *privately held* and *publicly expressed* opinions are studied using a novel model proposed by the author. There is a rich literature on such discrepancies. In [Waters and Hans, 2009], the authors recorded that over one third of jurors in criminal trials would have privately voted against the final decision of the jury panel they were on. The economist and political scientist Timur Kuran popularised the concept of *preference falsification*, where an individual presents a falsified view due to social pressure (be it imaginary or real) [Kuran, 1997]. Interestingly, there is often a disconnect between the true opinions of the general majority and what is perceived to be the true opinion of the general majority. This relates to the phenomenon of *pluralistic ignorance*; in the 1960s, the majority of white Americans were in fact *privately against segregation* but most individuals *assumed* there was a majority *in support of segregation* [O’Gorman, 1975]. One reason such discrepancies arise is due to a pressure to conform to a social norm, as illustrated by Solomon E. Asch’s seminal experiments on conformity [Asch, 1951]. The interested reader is referred to the introduction of Chapter 3, where these ideas are given a full treatment. The proposed model is used to revisit Asch’s experiments, and identify a situation where pluralistic ignorance arises in a network.

In the above literature, much focus was placed on modelling how the influence weights w_{ij} change as individuals interact. Beyond the simple assumption that $w_{ij}(t)$

changes as an explicit function of time t , the changes are often driven by a social process, e.g. homophily in the Hegselmann–Krause model or bias assimilation in [Dandekar et al., 2013]. Part II investigates how the w_{ij} change as captured by the DeGroot–Friedkin model [Jia et al., 2015]. The DeGroot–Friedkin model considers a social network which discusses opinions on a *sequence of topics*. At the end of each topic, individual i updates his/her self-weight w_{ii} (termed self-confidence or individual social power in [Jia et al., 2015]) via the social process of *reflected self-appraisal*. In particular, w_{ii} increases (respectively decreases) if individual i had a significant (respectively negligible) impact during the discussion of the previous topic. Owing to the assumption $\sum_{j=1}^n w_{ij} = 1$, as w_{ii} increases or decreases, necessarily w_{ij} for $j \neq i$ also changes. Using a variety of tools, investigations are conducted in Part II to identify the role of the network structure in determining how individual social power w_{ii} changes over the sequence of issues, and to draw conclusions on the limiting behaviour of w_{ij} . A modified variant of the DeGroot–Friedkin model was used to experimentally identify how social groups, over time, tend to become dominated by a single individual with high social power [Friedkin et al., 2016a].

Part III investigates a continuous-time model that describes how a network of individuals simultaneously discuss opinions on a set of logically interdependent topics. A key aspect of this work is that significant, as opposed to merely technical, differences in the conditions which guarantee that the opinion system remains stable are identified between the discrete-time model in [Parsegov et al., 2017; Friedkin et al., 2016b] (as noted earlier) and the continuous-time model proposed in this thesis. These differences are related to the network topology and matrix describing the logical interdependence.

1.2.4 Motivations and Key Concepts

Having introduced and discussed a number of different results, one question that requires answering is: why is there a need to develop agent-based opinion dynamics models? It is unlikely that such models will be able to capture human interactions in a social network to the same degree of accuracy as an Euler–Lagrange equation might capture the dynamics of a robotic manipulator arm [Spong et al., 2006], even though there are experimental validations of the DeGroot and Friedkin–Johnsen models as detailed earlier. In the author’s opinion, several reasons motivate the work.

First, a key objective is to be able to obtain high level observations about the conditions (both on the agent- and network-level) required to generate specific network-level phenomena, e.g. reaching a consensus. One may then use the model to explain how agent-level dynamics might, via networked interactions, create the observed social phenomena. This can be used to enrich the existing sociological and social psychological literature which typically study the same phenomena from the perspective of experiments or more qualitative methods. As an example, the goal of Part I of this thesis is precisely to re-examine interesting and established social phenomena not yet considered in agent-based models. At times, the agent-level dynamics can lead to surprising and unexpected results, such as when each individual’s desire to be

unique leads to overall conformity in the network [Smaldino and Epstein, 2015].

Second, well-posed and validated agent-based models can allow for a systematic determination of the key factors or parameters that govern the phenomena, and also allow for observations on how the phenomena may change as a function of the factors or parameters. This may not be straightforward for experimental or qualitative based methods at the beginning. As an example, and as will be explicitly detailed in Chapter 2, the DeGroot model reaches a consensus if the network satisfies certain connectivity requirements, while the final consensus value is in part determined by the individuals' "eigenvector centralities", which is a quantity dependent on the network structure and the strength of the interpersonal influence weights w_{ij} . This knowledge may prove crucial in identifying methods for reducing unwanted social phenomenon (reducing one type of unwanted social behaviour will be a major focus of Chapter 4). As a result, such models (provided they are experimentally validated later) will allow for *predictions to be made* on how changes to the network or to individuals' parameters may affect the observed social phenomena; this may also guide the design of future experimental and theoretical studies.

For the systems and control community to provide lasting contributions to this field, some key concepts should be kept in mind. The model must be sufficiently simple so that the full array of techniques developed in systems and control can be utilised. Some of the above works, e.g. [Mäs et al., 2014; Duggins, 2017] are able to capture a rich variety of phenomena but are so complex in the nonlinearities of the agent equations that there is little hope of obtaining analytical results or drawing qualitative or even general conclusions about the relation between behaviour and model parameters. (A parameter search might be employed to characterise the model's behaviour, but such models often have dozens of parameters, and so a curse of dimensionality exists). On the other hand, simple models may not be able to capture, as accurately, observed real-world social phenomena, which are often complex and unexpected. Thus, a difficult balance must be struck whereby the models must be sufficiently simple for analysis (including the drawing of qualitative or semi-quantitative conclusions) but with enough parameters and complexities to capture the phenomena of interest. Such a view is not unique to the author; an eloquent discussion is given in [Proskurnikov and Tempo, 2017]. Last, and to maximise the chances that the works are accepted beyond the systems and control community, one must ensure that the models proposed or studied are grounded in literature from sociology and social psychology, whether it is in the agent equations or the phenomena to be studied. This is a key difference from perhaps more conventional research on the closely related discipline of multi-agent systems, where control algorithms are designed to achieve a specific control objective. More is said in the next subsection.

1.2.5 Relation to Coordination of Multi-Agent Systems

The curious reader may be wondering why the systems and control community has displayed a recent and vigorous interest in problems on opinion dynamics. From one point of view, it is natural for systems and control researchers to seek to study com-

plex networked systems in a more inter-disciplinary setting. While this is true, the author's view is that opinion dynamics has been of particular interest due to its close relation to the problem of coordinating multi-agent systems, which has received sustained and high levels of interest over the past 2 decades. Specifically, the DeGroot and Abelson models are closely related to discrete-time and continuous-time algorithms, respectively, for coordinating groups of autonomous vehicles [Jadbabaie et al., 2003; Ren and Beard, 2005; Cao et al., 2008]. Also, the Taylor model was investigated as an algorithm for containment-control of a multi-agent system [Cao et al., 2012]. In more detail, opinion dynamics models have agent-level dynamics which lead to network-level social phenomena via interactions, while distributed algorithms are used to establish how autonomous agents interact with its neighbours in order to, e.g. reach a consensus [Ren and Beard, 2005; Olfati-Saber et al., 2007; Cai and Ishii, 2012], synchronise its states [Chopra, 2012; Wu et al., 2016; Qin and Yu, 2013] or form a geometric formation shape [Krick et al., 2009; Anderson et al., 2008].

The works reported in this thesis are therefore closely aligned with the author's other works conducted during his PhD studies on topics involving coordination of multi-agent systems. This includes a body of work on (i) consensus and synchronisations of networked Euler–Lagrange agents (listed as papers [6], [7], [8], [19], [20], [22], [23] on page vii), (ii) cooperative control of Global Positioning System (GPS)-denied Unmanned Aerial Vehicles (UAVs) using bearing measurements (listed as papers [3], [9], [18], [21], on page vii), and (iii) various other topics in multi-agent coordination (listed as papers [4], [12], [15], [17], [25], on page vii). One key difference is that in cooperative control of multi-agent systems, distributed control laws are developed for each agent to ensure the multi-agent system achieves some predefined objective, e.g. forming a geometric formation shape. Thus, the algorithms are limited only by the agents' capabilities (e.g. actuator saturation, or communication range). In opinion dynamics, the focus is on accurately modelling how an individual obtains and processes information, and establishing what type of behaviour may arise.

1.3 Thesis Outline and Statements of Collaborations

This thesis consists of ten chapters, including the current chapter. The next chapter introduces the fundamental DeGroot and Friedkin–Johnsen models in detail, and the concept of using a graph to capture a dynamic social network. Seven main chapters follow, each presenting main technical results. Chapter 10 provides conclusions and discusses future work. A short appendix gives a list of results on linear algebra, graph theory, and nonlinear contraction analysis used in the thesis. A brief outline of the thesis structure and the contents of each chapter is now presented. According to the different emphases of research topics and proposed research problems, the seven main chapters (Chapters 3-9) are divided into three parts as shown in the list. Figure 1.1 gives an overview of the thesis structure and the relation between the chapters and the established models discussed earlier.

After the summary of each main chapter (Chapters 3-9) given below, a brief sum-

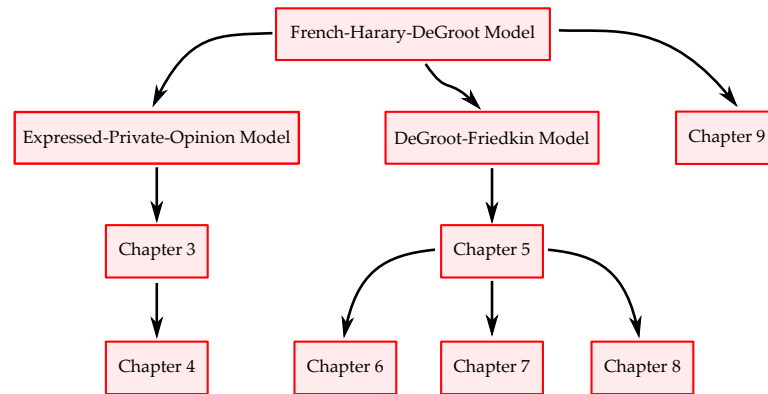


Figure 1.1: The structure of the thesis chapters and the models studied.

mary of the nature of the collaboration is reported when appropriate. The author is primarily responsible for the contribution of research outcomes in each main chapter.

1. Chapter 1 presents a general introduction to the research background, and covers some relevant models and results, as well as providing motivation for this thesis.
2. Chapter 2 introduces the DeGroot model of opinion dynamics, which is the fundamental model that this thesis builds upon. Standard convergence results and an interpretation of the social network as a graph are presented. The Friedkin–Johnsen model is also presented.

Part I: How Differences in Private and Expressed Opinions Arise

This part consists of Chapter 3 and Chapter 4.

3. Chapter 3 introduces a novel opinion dynamics model, termed the Expressed–Private–Opinion (EPO) model, which aims to show that a pressure to conform with the group norm can lead to an individual having different private and expressed opinions. Convergence results are obtained, and detailed analysis is conducted on the steady-state opinion configuration to draw conclusions within the social networks context.

This is joint work with Y. Qin, A. Govaert, and M. Cao (University of Groningen, Netherlands), and B. D. O. Anderson. Part of the research in this chapter was performed during a research visit to University of Groningen, Netherlands.

4. Chapter 4 uses the EPO model to examine and explain results from the seminal conformity experiments by Solomon E. Asch, which also provides a level of validation of the model. In addition, it is shown that the dangerous social phenomenon of “pluralistic ignorance” (explained in detail in this chapter) can occur due to the presence of stubborn extremists.

This is joint work with Y. Qin, A. Govaert, and M. Cao (University of Groningen, Netherlands), and B. D. O. Anderson. Part of the research in this chapter was performed during a research visit to University of Groningen, Netherlands.

Part II: Evolution of Individual Social Power

This part consists of Chapter 5, Chapter 6, Chapter 7 and Chapter 8.

5. Chapter 5 revisits the recently proposed DeGroot–Friedkin model of social power evolution over a sequence of issues. Nonlinear contraction analysis is employed to obtain an exponential convergence result, and other techniques are employed to gain insight into convergence rate and provide an upper bound on the social power of an individual at equilibrium.

This is joint work with J. Liu (SUNY at Stony Brook, USA), Tamer Başar (University of Illinois at Urbana-Champaign, USA), C. Yu and B. D. O. Anderson.

6. The main contribution of Chapter 6 is to extend the original DeGroot–Friedkin model to include dynamic (issue-varying) network topology. Nonlinear contraction analysis is again employed to draw the conclusion that individuals exponentially forget their initial social power; social power as the sequence of issues tends to infinity depends only on the network topology.

This is joint work with J. Liu (SUNY at Stony Brook, USA), Tamer Başar (University of Illinois at Urbana-Champaign, USA), C. Yu and B. D. O. Anderson.

7. Chapter 7 studies star graphs, which give rise to autocratic social power configurations. A number of network modification strategies are investigated, involving the introduction of new individuals and new interpersonal relationships in order to change the ordering of individual social power.

This is joint work with J. Liu (SUNY at Stony Brook, USA), Tamer Başar (University of Illinois at Urbana-Champaign, USA), C. Yu and B. D. O. Anderson.

8. Chapter 8 explores the original DeGroot–Friedkin model from the perspective of Lefschetz fixed-point theory. In particular, a simple result is obtained to conclude a local exponential convergence property, without the detailed calculations that arise when using nonlinear contraction analysis.

This is joint work with B. D. O. Anderson.

Part III: Opinion Dynamics with Interdependent Topics

This part consists of Chapter 9.

9. Chapter 9 considers a continuous-time model where individuals simultaneously discuss opinions on multiple logically interdependent topics. Necessary and sufficient conditions are established on the logical interdependency and network topology for ensuring a consensus of opinions is reached; it turns out that these conditions differ significantly from the discrete-time model. Stubborn individuals are also investigated.

The results in this chapter are from joint work with M. H. Trinh, and H.-S. Ahn (GIST, South Korea) and Y.-H. Lim (GNTECH, South Korea) and B. D. O. Anderson. Part of the research in this chapter was performed during a research visit to Gwangju Institute of Science and Technology, South Korea.

Preliminaries

This chapter introduces the notation to be used in the thesis, and an overview of graph theory. In addition, the DeGroot and Friedkin–Johnsen models are revisited.

2.1 Notations and Definitions

The notation used in this thesis is largely standard in systems and control literature. The set of real, integer, and natural numbers are denoted by \mathbb{R} , \mathbb{Z} , and \mathbb{N} , respectively. Unless stated otherwise, scalars and vectors are denoted by non-bold and bold lower case, respectively. That is, x is a scalar while \mathbf{x} is a vector. A matrix is denoted by a bold upper case variable, e.g. \mathbf{M} . Moreover, vectors are viewed as column vectors unless otherwise stated. The transpose of a vector \mathbf{x} or matrix \mathbf{M} is denoted by \mathbf{x}^\top or \mathbf{M}^\top , respectively. The i^{th} entry of \mathbf{x} and $(i, j)^{\text{th}}$ entry of \mathbf{M} are denoted by x_i and m_{ij} , respectively. The matrix $\text{diag}(x_i)$ denotes a diagonal matrix with the i^{th} entry being x_i .

For any real number x and complex number z , $|x|$ and $|z|$ denote the absolute value of x and modulus of z , respectively. The imaginary unit is denoted by $\sqrt{-1} = j$ and for a complex number $z = a + bj$, denote $\Re(z) = a$ and $\Im(z) = b$. The 1-norm, 2-norm, and ∞ -norm of a vector \mathbf{x} , and the associated induced matrix norm, are denoted by $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$, respectively. The n -column vector of all ones and zeros is given by $\mathbf{1}_n$ and $\mathbf{0}_n$, respectively. The $n \times n$ identity matrix is given by \mathbf{I}_n . The i^{th} canonical base unit vector of \mathbb{R}^n is denoted as \mathbf{e}_i , i.e. $\mathbf{e}_i \in \mathbb{R}^n$ has one in its i^{th} entry and zeros elsewhere. The Kronecker product is denoted by \otimes .

A matrix \mathbf{A} is said to be nonnegative (respectively positive) if all of its entries a_{ij} are nonnegative (respectively positive). The matrix \mathbf{A} is denoted as being nonnegative and positive by $\mathbf{A} \geq 0$ and $\mathbf{A} > 0$ respectively. Similarly, for a vector $\mathbf{x} \in \mathbb{R}^n$, $0 \leq \mathbf{x}$ and $0 < \mathbf{x}$ indicate component-wise inequalities, i.e., for all $i \in \{1, \dots, n\}$, $0 \leq x_i$ and $0 < x_i$, respectively. The n -simplex is $\Delta_n = \{\mathbf{x} \in \mathbb{R}^n : 0 \leq \mathbf{x}, \mathbf{1}_n^\top \mathbf{x} = 1\}$. Define $\tilde{\Delta}_n = \Delta_n \setminus \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ and $\text{int}(\Delta_n) = \{\mathbf{x} \in \mathbb{R}^n : 0 < \mathbf{x}, \mathbf{1}_n^\top \mathbf{x} = 1\}$ as the simplex excluding the corner points and the interior of the simplex, respectively.

A nonnegative matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$ is said to be row-substochastic (row-stochastic) if, for all i , there holds $\sum_{j=1}^m a_{ij} \leq 1$ ($\sum_{j=1}^m a_{ij} = 1$). A matrix \mathbf{A} is said to be doubly stochastic if $\sum_{j=1}^n a_{ij} = 1$ and $\sum_{i=1}^n a_{ji} = 1$. For two nonnegative matrices $\mathbf{A}, \mathbf{B} \geq 0$,

$A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{n \times m}$ are said to be of the same type, denoted by $A \sim B$, if A and B have strictly positive elements at the same positions. The spectral radius of a square matrix $A \in \mathbb{R}^{n \times n}$ is the largest modulus value of the eigenvalues of A , and is denoted by $\rho(A)$. For a matrix $A \in \mathbb{R}^{n \times n}$, $\lambda_i(A)$ denotes an eigenvalue of A . If $A = A^\top$, the eigenvalues are assumed to be ordered from smallest to largest, so that $\lambda_{\min}(A) \triangleq \lambda_1(A) \leq \dots \leq \lambda_n(A) \triangleq \lambda_{\max}(A)$. A matrix is irreducible if there does not exist a permutation matrix P such that $\bar{A} = P^\top A P$ has the lower triangular form

$$\bar{A} = \begin{bmatrix} \bar{A}_{11} & \mathbf{0} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix}, \quad (2.1)$$

where \bar{A}_{ij} have size greater than 0. A useful definition for a certain matrix property is now given.

Definition 2.1 (Primitivity, [Bullo et al., 2009, Definition 1.12]). *A nonnegative square matrix A is primitive if there exists $k \in \mathbb{N}$ such that $A^k > 0$.*

2.2 Graph Theory

As will become evident in the following section, a graph is a mathematical object that is a convenient and powerful method for representing the interactions between a group of individuals. The terms “network” and “graph” will be used interchangeably when there is no risk of confusion.

Given any square nonnegative matrix $A \in \mathbb{R}^{n \times n}$, one can associate with it a graph $\mathcal{G}[A] = (\mathcal{V}, \mathcal{E}[A], A)$. Here, $\mathcal{V} = \{v_1, \dots, v_n\}$ is the set of nodes of the graph $\mathcal{G}[A]$, with index set $\mathcal{I} = \{1, \dots, n\}$. An edge $e_{ij} = (v_i, v_j)$ is in the set of ordered edges $\mathcal{E}[A] \subseteq \mathcal{V} \times \mathcal{V}$ if and only if $a_{ji} > 0$. Unless specified, self-loops may exist, i.e. e_{ii} is allowed to be in \mathcal{E} , and thus $a_{ii} > 0$ is allowed. The edge e_{ij} is said to be incoming with respect to v_j and outgoing with respect to v_i , i.e. the arrow points¹ from v_i to v_j , and the weight a_{ji} is said to be the weight associated with edge e_{ij} . Fig. 2.1 gives an example network. The interpretation of the edges is dependent on the context and model, but typically connotes a form of directed interaction from individual i to individual j . In the opinion models considered in this thesis, an edge e_{ij} connotes that individual j learns of, and is in some way influenced by individual i 's opinion. Another interpretation is that the edge e_{ij} implies that individual i exerts an influence on individual j . The works considered in this thesis will only consider graphs with positive edge weights, i.e. $a_{ij} \geq 0$. Other works may consider negative edge weights to represent antagonistic or competitive interpersonal influence.

Since the graph is directed, in general, existence of e_{ij} does not imply existence of e_{ji} (and vice versa) and thus A is not assumed to be symmetric. The neighbour set of v_i is denoted by $\mathcal{N}_i = \{v_j \in \mathcal{V} : (v_j, v_i) \in \mathcal{E}\}$, and records the nodes j with edges

¹In some literature, and perhaps due to problem context, edge weights are defined so that $e_{ij} > 0 \Leftrightarrow a_{ij} > 0$. For a given $a_{ij} > 0$, the direction of the associated edge is therefore reversed from what is detailed in this thesis. The result is that the matrix A is unchanged, while all edges are drawn in the opposite direction. No issues arise in terms of analysis, other than use of different terminology.

incoming to i . Because self-loops are allowed, v_i may be in \mathcal{N}_i . A directed path is a sequence of edges of the form $(v_{p_1}, v_{p_2}), (v_{p_2}, v_{p_3}), \dots$, where $v_{p_i} \in \mathcal{V}$ and $e_{p_i p_{i+1}} \in \mathcal{E}$. Node i is reachable from node j if there exists a directed path from v_j to v_i . Moreover, a graph $\mathcal{G}[A]$ is strongly connected if and only if there is a path from every node to every other node [Godsil and Royle, 2001]. A graph $\mathcal{G}[A]$ is *strongly connected* if and only if A is *irreducible* [Godsil and Royle, 2001], or equivalently, there does not exist a reordering of the nodes \mathcal{V} such that A can be expressed in the form in Eq. (2.1). A directed cycle is a directed path that starts and ends at the same vertex, and contains no repeated vertex except the initial (which is also the final) vertex. The length of a cycle is the number of edges in the cyclic path. The periodicity of a directed graph is the smallest integer k that divides the length of every cycle of the graph. A graph is *aperiodic* if $k = 1$ [Bullo et al., 2009]. Note that any graph with a self-loop is aperiodic.

From results on nonnegative matrices and, further, the Perron-Frobenius Theorem (see Appendix A.1), one concludes that A has left and right eigenvectors with nonnegative entries, \mathbf{u}^\top and \mathbf{v} , associated with the eigenvalue $\lambda_1 = \rho(A)$. If $\mathcal{G}[A]$ is strongly connected, then $\lambda_1 = \rho(A)$ is a simple eigenvalue, \mathbf{u}^\top and \mathbf{v} can be taken to have strictly positive entries. All other eigenvalues $\lambda_i, i \neq 1$ satisfying $|\lambda_i| = \rho(A)$ are simple. There are precisely k eigenvalues of A (including $\lambda_1(A)$) with modulus equal to $\rho(A)$, where k is the periodicity of the graph $\mathcal{G}[A]$. If A is row-stochastic then $\lambda_1 = 1$, and the following definition is employed

Definition 2.2 (Dominant Eigenvectors). *For a strongly connected graph $\mathcal{G}[A]$, with row-stochastic A , let the strictly positive vector \mathbf{u}^\top and $\mathbf{1}_n$ be left and right eigenvectors of A associated with the simple eigenvalue $\lambda_1 = \rho(A) = 1$. With normalisation satisfying $\mathbf{u}^\top \mathbf{1}_n = 1$, \mathbf{u}^\top and $\mathbf{1}_n$ are termed the dominant left and right eigenvectors of A , respectively.*

Some literature refer to u_i as the eigenvector centrality of node i in the network [Newman, 2010] (other literature may not require A to be row-stochastic). In fact, eigenvector centrality will be a critical part of the DeGroot–Friedkin model studied in Part II. A result linking $\mathcal{G}[A]$ to the primitivity of A is now given.

Lemma 2.1 (Strongly connected, aperiodic graphs and primitive matrices, [Bullo et al., 2009, Proposition 1.35]). *The graph $\mathcal{G}[A]$ is strongly connected and aperiodic if and only if A is primitive, i.e. $\exists k \in \mathbb{N}$ such that A^k is a positive matrix.*

2.3 Basic Models of Opinion Dynamics

Two fundamental models of interpersonal influence networks are now reviewed, and basic convergence results stated. For clarity, the models are presented formally first in Sections 2.3.1 and 2.3.2, with comments and discussions on the model parameters and social context given in Section 2.3.3.

2.3.1 The DeGroot Model

Consider a population of n individuals, represented by $\mathcal{G}[W] = (\mathcal{V}, \mathcal{E}[W], W)$; the nodes are the individuals while the edges represent interpersonal influences. At time

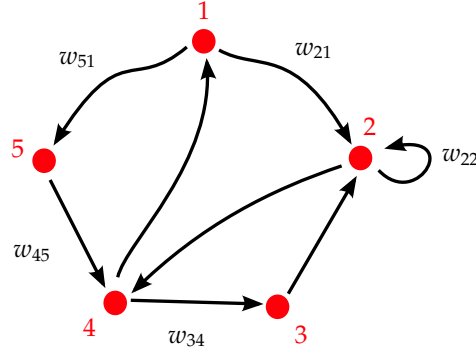


Figure 2.1: An example social network of 5 individuals modelled as the directed graph $\mathcal{G}[\mathbf{W}]$. The direction of the edges can be considered to represent influence flow. For example, individual 1 exerts a direct influence onto individual 2 to change 2's opinion via the directed edge e_{12} , with influence weight w_{21} . Individual 1 indirectly influences individual 3 via the directed path formed by the sequence of edges e_{15}, e_{54}, e_{43} , with weights w_{51}, w_{45}, w_{34} .

steps $0, 1, \dots$, individual i 's opinion² $y_i(t) \in \mathbb{R}$ evolves according to following:

$$y_i(t+1) = \sum_{j=1}^n w_{ij} y_j(t). \quad (2.2)$$

With $w_{ij} \geq 0$, it is assumed that for all $i \in \mathcal{I}$, there holds $\sum_{j=1}^n w_{ij} = 1$, which implies that \mathbf{W} is row-stochastic. In some literature, the notation $\sum_{j \in \mathcal{N}_i} w_{ij} y_j(t)$ is used; this is equivalent to the right hand side of Eq. (2.2) because for any $j \notin \mathcal{N}_i$, $w_{ij} = 0$. It should be noted that all individuals update their opinions synchronously; asynchronous models are available but will not be considered further in this thesis. For convenience, the vector of opinions $\mathbf{y}(t) = [y_1(t), \dots, y_n(t)]^\top$ is used to represent the opinions of all individuals in the network. The compact form of the opinion dynamical system for the influence network $\mathcal{G}[\mathbf{W}]$ with each individual's opinion evolving according to Eq. (2.2), is given by

$$\mathbf{y}(t+1) = \mathbf{W}\mathbf{y}(t). \quad (2.3)$$

The matrix \mathbf{W} is sometimes referred to as the influence matrix. The following is a standard definition for the final opinion distribution.

Definition 2.3. Let $\mathbf{y}(t) = [y_1(t), \dots, y_n(t)]^\top$ be the vector of the opinions of individuals $1, \dots, n$ in a social network. The opinions are said to have reached a consensus if and only if $\mathbf{y}^* \triangleq \lim_{t \rightarrow \infty} \mathbf{y}(t)$ satisfies $\mathbf{y}^* = \alpha \mathbf{1}_n$, for some finite $\alpha \in \mathbb{R}$.

A standard result on the convergence of the DeGroot model to consensus is now given, see e.g. [Ren and Beard, 2007; Proskurnikov and Tempo, 2017; Bullo et al., 2009; Nedić and Liu, 2017].

²See Section 2.3.3 for a discussion how an opinion can be represented as a real number.

Theorem 2.1. *Suppose that the opinion $y_i(t)$ of each individual i in the network $\mathcal{G}[\mathbf{W}]$ evolves according to Eq. (2.2). Suppose further that $\mathcal{G}[\mathbf{W}]$ is strongly connected and aperiodic, and \mathbf{W} is row-stochastic. Then, $\lim_{t \rightarrow \infty} \mathbf{y}(t) = (\boldsymbol{\zeta}^\top \mathbf{y}(0)) \mathbf{1}_n$ exponentially fast, where $\boldsymbol{\zeta}^\top$ is the dominant left eigenvector of \mathbf{W} .*

The author would like to point out that, owing to the problem context, most strongly connected social networks are also aperiodic. This is because it is rare to have a network where there does not exist at least one individual i with a nonzero self-weight w_{ii} , which for strongly connected networks immediately implies aperiodicity. Interestingly, while Theorem 2.1 gives a result for exponential convergence to a consensus, it is in fact possible for Eq. (2.3) to achieve finite-time consensus for certain types of network topologies [Hendrickx et al., 2015]. That is, $\mathbf{y}(t) = (\boldsymbol{\zeta}^\top \mathbf{y}(0)) \mathbf{1}_n$ for all $t \geq T_1$, with $T_1 < \infty$. In fact, the set of topologies which achieves finite-time consensus appears to not be a thin set. This is in stark contrast to the continuous-time Abelson model, which may also reach a consensus exponentially fast but never in finite time. Convergence results do exist for $\mathcal{G}[\mathbf{W}]$ which are not strongly connected, e.g. which contain a directed spanning tree (see Appendix A.2 for a definition), but these will not be examined, since Parts I and II deal with strongly connected graphs. Part III does consider graphs that are not strongly connected, but the model is in continuous-time, and relevant results are deferred to that part. A semi-contraction result is now given (the proof is straightforward, see e.g. [Proskurnikov and Tempo, 2017]).

Corollary 2.1. *Suppose that the opinion $y_i(t)$ of each individual i in the network $\mathcal{G}[\mathbf{W}]$ evolves according to Eq. (2.2), where \mathbf{W} is row-stochastic. Define $\bar{y}(t) = \max_{i \in \mathcal{I}} y_i(t)$ and $\underline{y}(t) = \min_{i \in \mathcal{I}} y_i(t)$. Then, there holds*

$$\bar{y}(t) \geq \bar{y}(t+1) \tag{2.4a}$$

$$\underline{y}(t) \leq \underline{y}(t+1) \tag{2.4b}$$

for all $t \geq 0$.

From this result, it is said that the system Eq. (2.3) is semi-contracting because $V(t) = \bar{y}(t) - \underline{y}(t)$ is nonincreasing over t . In the problem context, this means that the opinions grow closer together in value or stay the same distance, but never become further apart in value.

The introduction to the DeGroot model is completed by noting, as discussed in Chapter 1, that generalisations exist which consider the influence weights as a function of time or of opinions $\mathbf{y}(t)$, or of both, i.e. $w_{ij}(t, \mathbf{y}(t))$. For clarity, in this thesis time-varying refers exclusively to influence weights that are dependent explicitly on time t , i.e. $w_{ij}(t)$, whereas the notation $w_{ij}(\mathbf{y}(t))$ is used to refer to a state-dependent influence weight. The focus of this thesis is to study new complex phenomena, as opposed to detailing convergence analysis with time-varying or state-dependent influence weights, and unless specifically stated otherwise, it is assumed that the w_{ij} are time-invariant and state-independent.

2.3.2 The Friedkin–Johnsen Model

As discussed in Chapter 1, there is interest in investigating how strong diversity arises in social networks (see Chapter 1 for a definition of strong diversity). In particular, Theorem 2.1 indicates that in strongly connected and aperiodic influence networks, social influence acts until opinions reach a consensus. Since many social networks are strongly connected, or contain a strongly connected component, one must ask “if social influence acts to bring opinions to a consensus what other process generates strong diversity?” To explain this, Noah Friedkin and Eugene Johnsen proposed a new model, known now as the Friedkin–Johnsen model, where the opinion $y_i(t)$ of individual i , in $\mathcal{G}[\mathbf{W}]$, evolves as

$$y_i(t+1) = \lambda_i \sum_{j=1}^n w_{ij} y_j(t) + (1 - \lambda_i) y_i(0). \quad (2.5)$$

Similar to the DeGroot model, it is assumed that $w_{ij} \geq 0$ and $\sum_{j=1}^n w_{ij} = 1$ for all i . Here, $\lambda_i \in [0, 1]$ represents individual i ’s *susceptibility or openness* to interpersonal influence, as captured by the term $\sum_{j=1}^n w_{ij} y_j(t)$. The term $1 - \lambda_i$ is sometimes referred to as individual i ’s *stubbornness*, as it represents attachment to his/her initial opinion $y_i(0)$. Thus, $\lambda_i = 1$ represents maximal openness to interpersonal influence as in the DeGroot model, while $\lambda_i = 0$ represents maximal closure to interpersonal influence. Some literature refer to individuals with $\lambda_i < 1$ as *prejudiced* individuals. It is clear that $y_i(t+1)$ is a convex combination of individual i ’s interpersonal influence $\sum_{j=1}^n w_{ij} y_j(t)$ and his/her initial opinion $y_i(0)$. (A natural and frequently assumed choice is $\lambda_i = 1 - w_{ii}$ [Friedkin and Johnsen, 1990; Friedkin and Bullo, 2017; Friedkin et al., 2016b]). For arbitrary λ_i , one obtains the compact form of the opinion dynamical system as

$$\mathbf{y}(t+1) = \mathbf{\Lambda} \mathbf{W} \mathbf{y}(t) + (\mathbf{I}_n - \mathbf{\Lambda}) \mathbf{y}(0), \quad (2.6)$$

where $\mathbf{\Lambda} = \text{diag}(\lambda_i)$, $i = 1, \dots, n$. Notice that if all $\lambda_i = 1$, then Eq. (2.6) reduces to Eq. (2.3). I.e., the DeGroot model is a special case of the Friedkin–Johnsen model where all individuals are maximally open to interpersonal influence. The following is a recent result from [Parsegov et al., 2017], rephrased for clarity in the context of this thesis.

Theorem 2.2 (Stability of the Friedkin–Johnsen Model). *Suppose that the opinion $y_i(t)$ of each individual i in $\mathcal{G}[\mathbf{W}]$ evolves according to Eq. (2.5). Suppose further that $\mathcal{G}[\mathbf{W}]$ is strongly connected, \mathbf{W} is row-stochastic, and there exists an $i \in \mathcal{I}$ such that $\lambda_i < 1$. Then,*

$$\mathbf{y}^* \triangleq \lim_{t \rightarrow \infty} \mathbf{y}(t) = (\mathbf{I}_n - \mathbf{\Lambda} \mathbf{W})^{-1} (\mathbf{I}_n - \mathbf{\Lambda}) \mathbf{y}(0), \quad (2.7)$$

with $\rho(\mathbf{\Lambda} \mathbf{W}) < 1$.

Again, it is noted that [Parsegov et al., 2017] does consider graphs which are not strongly connected, but the results are omitted here. Interestingly, while the DeGroot

model has extensive results on time-varying networks $\mathcal{G}[\mathbf{W}]$, results on time-varying networks under the Friedkin–Johnsen model have only appeared recently [Proskurnikov et al., 2017].

Note that for strongly connected $\mathcal{G}[\mathbf{W}]$ where every individual has some stubbornness, i.e. $\lambda_i < 1$ for all $i \in \mathcal{I}$, there holds $y_i^* \neq y_j^*, \forall i, j$, generically. That is, no individuals have equal opinion values at equilibrium, which implies strong diversity among the final opinion values. To the author’s knowledge, this last result is not available in the literature, but can be derived following a method similar to what will be presented in Section 3.3. Also of note is that the property set out in Corollary 2.1 does not hold in general for the Friedkin–Johnsen model. However, one can show that for all $i \in \mathcal{I}$ and $t \geq 0$, $y_i(t) \in [a, b]$ where $a = \min_{i \in \mathcal{I}} y_i(0)$ and $b = \max_{i \in \mathcal{I}} y_i(0)$. Again, these results can be proved using methods detailed in Section 3.3.

2.3.3 Comments on the Models

Having presented the models formally as dynamical systems, some brief commentary is now given on the models in light of their context in influence network modelling. These comments should be considered throughout the thesis, as they will be relevant to all works presented.

Representation of an Opinion: First, the representation of an opinion $y_i(t)$ as a real number is elucidated. Several applications exist in which the definition of y_i as a real number is useful. For example, the social network may be discussing a topic which is defined by a question with a necessarily subjective answer, e.g. “is pasta tasty?” Then, negative and positive values of y_i represent disagreement and agreement, respectively, while $y_i = 0$ represents a neutral stance. The magnitude of $y_i(t)$ represents the intensity of the agreement or disagreement. Alternatively, one could consider the topic as a statement on an idea, e.g. “same-sex marriage should be legalised”, with y_i representing individual i ’s attitude towards the idea. Then, negative and positive values represent i opposing and supporting the idea, respectively. Depending on the model and problem context, it may be useful to scale the opinions so that $y_i(0) \in [-1, 1]$, with $y_i = -1$ and $y_i = 1$ representing the extreme opinions on the spectrum; a well-defined model (such as the DeGroot and Friedkin–Johnsen models) would then have the property that $y_i(t) \in [-1, 1]$ for all $t \geq 0$. Yet other works consider $y_i(t) \in [0, 1]$. Some works, e.g. [Yildiz et al., 2013; Nowak et al., 1990], consider $y_i(t)$ as a discrete variable (with one typical choice being binary 0, 1). These may be suitable for opinions that lead to actions being taken, e.g. y_i may represent the voting choice for individual i in a political election, or a choice on whether to buy the latest smart phone. This thesis elects to consider $y_i \in \mathbb{R}$ as it better captures *differences* in opinions, such as the differences that can arise between an individual’s private and expressed opinion, as studied in Chapter 3 and 4.

Terminology: The terms “opinion”, “attitude”, “belief” are just a few of many that appear in the social science literature. There are some subtle differences, with distinctions made difficult due to a lack of consistent and agreed upon definitions across different scientific communities. In this thesis, the author takes the view that an indi-

vidual's *belief* is his/her position on a statement which is provable to be true or false, e.g. "the Earth orbits around the Sun." An individual's *opinion* is his/her position on a subjective statement which cannot be proved to be true or false, e.g. "vanilla ice cream tastes better than chocolate ice cream." This distinction is simply one choice of the definitions from the many possible versions in the literature. No further attempt is made to distinguish the terms, and unless stated otherwise, this thesis will exclusively use the term "opinion" when referring to $y_i(t)$. This is because the models considered in this thesis are general enough to cover many scenario applications.

Multiple Topics: If the individuals are discussing m independent topics, then one can define $\mathbf{y}_i(t) = [y_i^1(t), \dots, y_i^m(t)]^\top$ as individual i 's vector of opinions, with y_i^k denoting i 's opinion on topic $k \in \{1, \dots, m\}$. The Kronecker product is used to trivially extend existing results, e.g. Theorem 2.1 and 2.2. For example, the DeGroot compact form becomes

$$\mathbf{y} = (\mathbf{W} \otimes \mathbf{I}_m) \mathbf{y}(t), \quad (2.8)$$

where $\mathbf{y}(t) = [\mathbf{y}_1^\top, \dots, \mathbf{y}_n^\top]^\top$. When the m topics become dependent on each other, new analysis is required, and Chapter 9 investigates one model's method of capturing interdependence among the topics.

Interpretation of Parameters: For parameters w_{ij} , λ_i and the other parameters that will be introduced in later models, it is clear that their values depend on many factors, such as individual i 's personality, culture, upbringing and experience, or whether i is an expert on the topic of discussion. The magnitudes of w_{ij} can depend on level of friendship, status and rank (formal or informal) of the individuals in the network, etc. This thesis does not aim to identify these values for a given social network, or explain how or why the parameters may be different for different individuals. The works in thesis only postulate that the parameters exist, and that the individuals' opinions evolve according to the models that will be later presented. It is not even assumed that individuals necessarily *know the exact value* themselves, or are aware that their opinions evolve as captured in the models. The key focus of this thesis is to consider *how* the opinions evolve for a given set of parameters, and draw *quantitative or semi-quantitative conclusions* on the effects of the parameters on the opinion evolution, which may be used to gain high-level insight into interpersonal influence networks.

Time-Scales: Last, it should be noted that the above models are typically suited for application on problems with short time-scales, e.g. a boardroom discussion lasting several hours or perhaps at a workshop over a week. Such models may not accurately reflect discussion over months or years, because almost certainly w_{ij} , λ_i , etc., would change over time (the precise nature of the time-variation depends on many factors). Chapters 5 through 8 do consider a model where the network discusses a sequence of issues, which may be appropriate for longer time-scales. However, that model assumes that each individual's self-weight w_{ii} changes after discussion on a issue, following a social process called reflected self-appraisal. Thus, the interest is in the evolution of w_{ii} over the issue sequence, and y_i is not the primary variable of interest.

Part I

How Differences in Private and Expressed Opinions Arise

A Novel Model for Opinion Dynamics Under Pressure to Conform

Part Summary

Part I studies a novel opinion dynamics model proposed by the author, termed the Expressed–Private–Opinion (EPO) model. The model draws inspiration from some of the most classical results, and aims to develop a mathematical framework for describing social phenomena involving individuals who have different expressed and private opinions. In the EPO model, each individual is assumed to have a private and expressed opinion, and each individual’s expressed opinion is altered from the individual’s private opinion by a pressure to conform to the social norm. Chapter 3 introduces the EPO model and presents theoretical results on stability, and analyses the private and expressed opinions at equilibrium. Chapter 4 uses the model to revisit the seminal conformity experiments by Solomon Asch and investigate how pluralistic ignorance can arise due to the presence of stubborn extremists.

Chapter Summary

This chapter introduces the Expressed–Private–Opinion (EPO) model. A number of phenomena involving individuals with different private and expressed opinions have been recorded and studied in the social sciences via qualitative and experimental methods, and these are discussed in the introduction immediately below. After introducing and explaining the motivation of the model, conditions on parameters of the influence network are obtained which guarantee that the opinions convergence to a steady state. Then, investigations are conducted to draw several interesting conclusions on the effects of individual stubbornness, and resilience to the pressure to conform to a group majority, in generating differences in an individual’s expressed and private opinion.

3.1 Introduction

In much of the existing literature on agent-based opinion dynamics modelling, it is assumed that each individual has an opinion $y_i(t)$ which is communicated to others in the network. Few models consider the possibility that an individual expresses an opinion different to his/her private opinion, even though the reader will almost certainly have been in a situation where this has occurred to them.

On the other hand, these situations are well studied in the social and political sciences. In [Waters and Hans, 2009], the authors found that over one third of jurors in criminal trials would have privately *voted against* the decision of the jury panel they were on. Large differences between the private and expressed opinions of the civilian population can generate discontent and tension, which might result in *violent and unforeseen actions* such as the Arab Spring movement [Goodwin, 2011] and the fall of the Soviet Union [Kuran, 1989]. During the rise of Islamic State in 2014, a US led coalition readily expressed agreement to attack Raqqa, then the de-facto capital of Islamic State. Later, when deciding which troops were to lead the ground assault, the *opposing private opinions* of the Turkish and Kurdish representatives in the coalition emerged and created a deadlock for almost two years [Mintz and Wayne, 2016]. Access to the public actions of individuals, without being able to observe their thought processes that led to the actions, can spark an informational cascade where *all successive individuals choose the wrong action* [Bikhchandani et al., 1992]. This was used to help explain why farmers in Iowa refused to adopt hybrid seed corn for years, despite its benefits [Ryan and Gross, 1943]. Due to fears of social isolation and exposure, some individuals enforce social norms despite privately disliking the norms [Centola et al., 2005; Willer et al., 2009].

Naturally, there is interest in identifying what creates such differences or discrepancies between expressed and private opinions/actions. One commonly hypothesised reason is that such differences arise due to a *pressure to conform to a group standard or norm*. Formal study of such pressure goes back many decades. In 1951, Solomon E. Asch's seminal paper [Asch, 1951] showed that individuals could react differently when their judgment about an indisputable fact was challenged by a unanimous majority. Some individuals could withstand the pressure, whereas the actions and judgments of other individuals were heavily affected. A variety of other studies have been reported, and they generally establish that such pressures not only generate different expressed and private opinions/actions, but can also have other consequences. In some instances, high productivity factory workers were pressured to lower their production rate to match factory averages [Coch and French Jr, 1948]. Peer punishment is often dealt to individuals who deviate from group norms, such as in gangs [Thrasher, 1963]. This occurs even if the norm is destructive or unhealthy for the group itself [Abbink et al., 2017]. The pressure exerted on an individual to conform may change over time, or depend on his/her opinions and/or the opinions/actions of others in the group [Waters and Hans, 2009; Asch, 1951; Schachter, 1951; Gorden, 1952].

Significant advances have been achieved, separately, in the model complexity and

analysis of the key factors in each model that determine specific dynamical properties (see Chapter 1). Despite this, existing agent-based models have failed to provide a thorough account of phenomena involving (i) differences in private and expressed opinions, and (ii) the effects of a pressure to conform to the group norm (both of which have been well-studied in the social and social psychology literature). There is therefore significant need and motivation to examine these interesting social phenomena from the perspective of agent-based models, and investigate the precise mechanisms that drive said phenomena. The aim is to, for the first time, provide a mathematical framework for the study of opinion evolution under pressure to conform. This chapter will introduce the EPO model, drawing inspiration from the established Friedkin–Johnsen model. In the proposed EPO model, each individual has both an expressed and a private opinion, and the expressed opinion is altered from the private opinion due to a pressure to conform; this is a key departure from most existing works. This chapter will focus on development and analysis of the model from a systems and control perspective, including the establishing of convergence results and drawing of semi-quantitative conclusions that give insight into how stubbornness and resilience to pressure to conform affect the expressed and private opinions of individuals in an influence network. In the subsequent Chapter 4, Solomon E. Asch’s seminal experiments are revisited using the proposed model, and the well-studied phenomenon of *pluralistic ignorance* is explained using the model.

3.1.1 Chapter Organization

The rest of this chapter is organized as follows. Section 3.2 introduces the opinion dynamics model. Section 3.3 provides convergence results and also establishes several semi-quantitative conclusions on the effects of individuals’ parameters on the final opinion distribution. Concluding remarks are given in Section 3.4

3.2 A Model with Expressed and Private Opinions

To begin, the opinion dynamics model is formally introduced. Then, explanations are provided on the motivation for modelling the evolution of the private and expressed opinions in the manner described. Comments are also given to clarify the model and its relation to existing works. To conclude, the model is presented in a compact form to aid the theoretical analysis that will occur in Section 3.3.

3.2.1 The Opinion Dynamics Model

For a population of n individuals, whose interactions are captured by a directed network $\mathcal{G}[\mathbf{W}]$, let $y_i(t)$ and $\hat{y}_i(t)$, $i \in \mathcal{I}$, represent, at time $t = 0, 1, \dots$, individual i ’s privately held and publicly expressed opinions on a given topic, respectively. In general, $y_i(t)$ and $\hat{y}_i(t)$ are not the same, and y_i is regarded as individual i ’s true opinion. Individual i may refrain from expressing $y_i(t)$ for many reasons, e.g. political

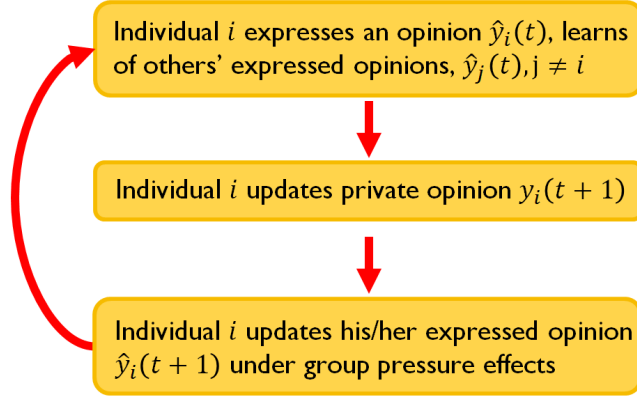


Figure 3.1: Each individual undergoes this process of influence and opinion evolution, occurring simultaneously with all other individuals. At time step t , individual i expresses opinion $\hat{y}_i(t)$ and learns of i 's neighbours' expressed opinions $\hat{y}_j(t), j \in \mathcal{N}_i$. This might occur in a boardroom meeting or in an online social medium. Next, individual i 's privately held opinion evolves to be $y_i(t+1)$ under the influence of i 's own privately held opinion $y_i(t)$ and the expressed opinions of i 's neighbours, $\hat{y}_j(t)$. The precise algorithm is given in Eq. (3.1). Once individual i has updated his/her privately held opinion, i then determines the new $\hat{y}_i(t+1)$ to be expressed in the next round of discussion, as given in Eq. (3.2).

correctness when discussing a sensitive topic, to avoid conflict, or to leave a favourable impression on another individual [Kuran, 1997] (see Chapter 4 for comments on preference falsification). Fig. 3.1 illustrates, and qualitatively describes, “the opinion dynamics process”.

Formally, individual i 's private opinion evolves as

$$y_i(t+1) = \lambda_i w_{ii} y_i(t) + \lambda_i \sum_{j \neq i}^n w_{ij} \hat{y}_j(t) + (1 - \lambda_i) y_i(0), \quad (3.1)$$

and individual i 's expressed opinion is determined according to

$$\hat{y}_i(t) = \phi_i y_i(t) + (1 - \phi_i) \hat{y}_{\text{avg}}(t-1). \quad (3.2)$$

The influence weight $w_{ij} \geq 0$ represents the weight individual i accords to his/her neighbour individual j 's expressed opinion $\hat{y}_j(t)$. As in the DeGroot and Friedkin–Johnsen models (see Chapter 2) it is assumed that for all $i \in \mathcal{I}$, $\sum_{j=1}^n w_{ij} = 1$. Similarly to the Friedkin–Johnsen model (Section 2.3.2), the constant $\lambda_i \in [0, 1]$ represents individual i 's susceptibility to interpersonal influence changing i 's private opinion ($1 - \lambda_i$ is i 's stubbornness in remaining attached to initial opinion $y_i(0)$). The quantity $\hat{y}_{\text{avg}}(t) = \frac{1}{n} \sum_{j=1}^n \hat{y}_j(t)$ is termed the *public opinion* of the influence network, and represents the evolving group standard or norm. The constant $\phi_i \in [0, 1]$ encodes individual i 's resilience to the pressures to conform to the public opinion (maximally $\phi_i = 1$, and minimally $\phi_i = 0$), and is referred to as i 's *resilience* for short. Observe

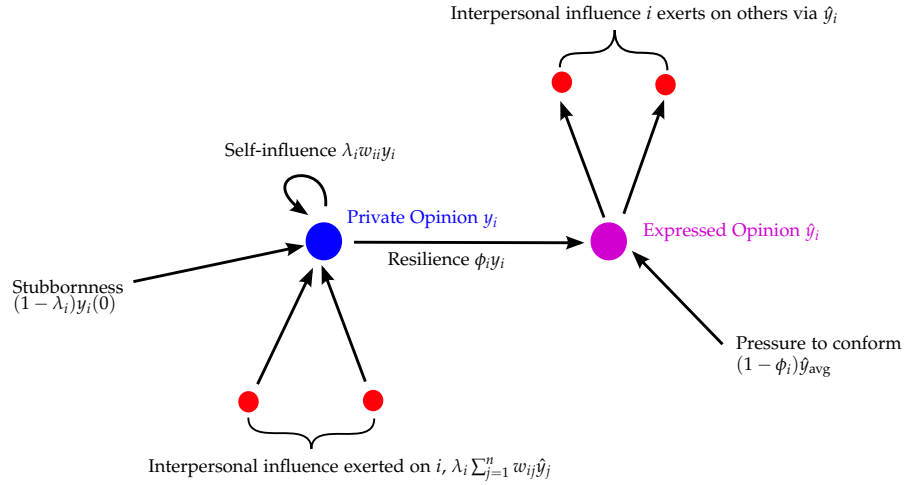


Figure 3.2: Individual i 's private opinion $y_i(t)$ and expressed opinion $\hat{y}_i(t)$ are represented by the blue node and purple node, respectively. The arrows/edges represent quantities which exert an influence on i 's private or expressed opinion. In particular, $y_i(t)$ is changed due to interpersonal influence from other individuals in the influence network, i 's self-weight (if there is a self-loop), and an attachment to initial opinion $y_i(0)$. The private opinion of individual i exerts an influence on i 's expressed opinion via a resilience to conformity, while the public opinion \hat{y}_{avg} exerts a pressure to conform. Individual i exerts an influence on others in the network via i 's expressed opinion \hat{y}_i .

that substituting $\hat{y}_j(t)$ from Eq. (3.2) into Eq. (3.1) yields

$$y_i(t+1) = \lambda_i w_{ii} y_i(t) + \lambda_i \sum_{j \neq i}^n w_{ij} \phi_j y_j(t) + \frac{\lambda_i}{n} \sum_{j \neq i}^n w_{ij} (1 - \phi_j) \mathbf{1}_n^\top \hat{\mathbf{y}}(t-1) + (1 - \lambda_i) y_i(0) \quad (3.3)$$

This gives a clear indication of the different terms which exert an influence to change i 's private opinion. The initial expressed opinion is set to be equal to the initial private opinion, i.e. $\hat{y}_i(0) = y_i(0)$, which means Eq. (3.1) comes into effect for $t = 1$. The influences that act on individual i to change i 's private and expressed opinion are illustrated in Fig. 3.2.

Notice that setting $\phi_i = 1$ for all i , i.e. all individuals express their private opinion unaltered, recovers the Friedkin–Johnsen model. Setting $\phi_i = \lambda_i = 1$, for all i , recovers the DeGroot model. From one perspective, this implies that the EPO model is a generalisation of both the Friedkin–Johnsen and DeGroot models. From a different perspective, this gives some support to the EPO model, since it builds upon, and is inspired by two established models which have been empirically validated (see Chapter 1). Further validation of the EPO model is provided in Chapter 4, where the model is used to examine Asch's conformity experiments, including to show that the experimental results are fully captured by the proposed model, and identify how the accepted social phenomenon of *pluralistic ignorance* can arise.

Model Explanation

Eq. (3.2) indicates that individual i 's expressed opinion $\hat{y}_i(t)$ is a convex combination of i 's private opinion, $y_i(t)$, and the public opinion from the previous time step, $\hat{y}_{\text{avg}}(t-1)$. The papers [Gorden, 1952; Asch, 1951] revealed that, perhaps unsurprisingly, normative pressure to conform was always exerted on an individual's opinion so that his/her expressed opinion moved from his/her private opinion *in the direction of the group norm* (in this case, the public opinion). In other words, Eq. (3.2) represents how individual i 's expressed opinion is his/her private opinion *modified or altered due to normative pressure*, proportional to $1 - \phi_i$, to be closer to the public opinion. Thus, the pressure to conform exerts a "force" $(1 - \phi_i)\hat{y}_{\text{avg}}(t-1)$ onto individual i to express an opinion in line with the group norm.

In that sense, $\phi_i = 1$ captures an individual i who is maximally resilient, and expresses precisely his/her private opinion $y_i(t)$, with no fear of social isolation or being viewed differently. An individual i with $\phi_i = 0$ is minimally resilient, and is someone who totally fears the appearance of being different from the group, and says exactly what the group's current view is. Some pressures of conformity may derive from unspoken traditions [Merei, 1949], and others arise because of desire to be in the group, driven by e.g. monetary incentives, status or rewards [Festinger, 1950]. As a result, heterogeneous ϕ_i are used to capture the fact that some individuals are less inhibited or reserved than others when expressing their opinions. In addition, pressures are exerted (or perceived to be exerted) differently for individuals, e.g. due to status or rank within a group [Schachter, 1951; Gorden, 1952].

Comments on the Model

A key feature in the EPO model is the defining of two states y_i, \hat{y}_i for each individual and the restriction that only other \hat{y}_j (and no y_j) may be available to individual i . A second key feature is the proposal of Eq. (3.2) as the model for capturing how an individual determines his/her expressed opinion, with motivation drawn from social psychology literature. As it turns out, these two features allow the model to capture phenomena that other agent-based opinion dynamics models are not able to capture. In particular, this model departs from many established models, some of which were covered in Chapter 1, and which assume only one opinion variable per individual e.g. [DeGroot, 1974; Hegselmann and Krause, 2002; Dandekar et al., 2013; Mäs et al., 2014; Friedkin and Johnsen, 1990].

The recent paper [Duggins, 2017] does have a model in which there are two states y_i, \hat{y}_i to represent the private and expressed opinion for each individual. However, in addition to being highly complex and nonlinear, that model is fundamentally different in how $y_i(t)$ and $\hat{y}_i(t)$ evolve. Many existing models e.g. [Duggins, 2017; Mäs et al., 2014; Nowak et al., 1990], are sufficiently complex that there are no straightforward methods to systematically determine *how the parameters affect the dynamical behaviour of the opinions*. It may be possible to conduct an exhaustive parameter sweep, but because there may be many parameters, a curse of dimensionality exists.

On the other hand, and by design, the EPO model attempts to strike a balance

so that it is simple enough for theoretical analysis to be conducted, but sophisticated enough to capture a range of social phenomena previously missing from agent-based models. This leads to multiple approaches of analysis which complement each other; simulations are used to Chapter 4 to study complex social phenomena and theoretical analysis is conducted in this chapter to gain insight into how certain parameters of individuals affect the behaviour of the overall network.

Remark 3.1. One can immediately notice the time-shift of $\hat{y}_{avg}(t-1)$ in Eq. (3.2). The time-shift is required because otherwise use of $\hat{y}_{avg}(t)$ would result in both the left and right hand side of Eq. (3.2) being dependent on $\hat{y}_i(t)$, which does not make sense and would create an inconsistent equation. Eq. (3.2) is also consistent with the process described in Figure 3.1. It will be shown in the sequel that such a time-shift can be dealt with. It was noted earlier that the initial conditions are assumed to be $\hat{y}_i(0) = y_i(0)$. As it turns out, under mild assumptions on the stubbornness of the individuals, the initial conditions $\hat{y}_i(0)$ for all $i \in \mathcal{I}$, are forgotten exponentially fast. Specifically, the equilibrium opinion values are dependent on the initial private opinions $y_i(0)$ but independent of the initial expressed opinions $\hat{y}_i(0)$. As a consequence one could initialise $\hat{y}_i(0)$ arbitrarily, and the final equilibrium opinions would be unchanged (though the transient would change)

Remark 3.2. One may ask: why does the new expressed opinion $\hat{y}_i(t+1)$ depend explicitly on the public opinion $\hat{y}_{avg}(t)$, but the new private opinion $y_i(t+1)$ does not depend explicitly on the public opinion? First, Eq. (3.3) indicates that the public opinion does in fact influence the update of individual i 's private opinion, reflected in the term $\mathbf{1}_n^\top \hat{\mathbf{y}}(t-1)/n$. This influence arises indirectly through neighbour j 's expressed opinion. It is also possible to adjust the model to include a direct influence by the public opinion in shaping the update of individual i 's private opinion. This is done by introduction of a “virtual agent”. Specifically an additional node is inserted into the influence network, labelled v_{n+1} . One can set $\phi_{n+1} = 1$ and the private opinion update is designed (easily) to ensure that $y_{n+1}(t) = \frac{1}{n} \sum_{j=1}^n \hat{y}_j(t), \forall t$. One can capture individual i 's private opinion being directly influenced by the public opinion by setting the influence weight $w_{i,n+1} > 0$. In relation to the theoretical results presented later in Section 3.3, one can verify that Theorems 3.1 and 3.2, and Corollary 3.1 will hold with minor adjustments. Corollary 3.3 also holds with minor adjustment. However, Corollary 3.2 will not hold.

3.2.2 Local Public Opinion

In small networks, e.g. a boardroom of 10 people, the public opinion \hat{y}_{avg} , is likely to be discernible to every individual. For large networks, such information may come from opinion polls or social media trends. As an alternative, it is also possible to study the model in which Eq. (3.2) is replaced by

$$\hat{y}_i(t) = \phi_i y_i(t) + (1 - \phi_i) \hat{y}_{i, \text{lavg}}(t-1) \quad (3.4)$$

where $\hat{y}_{i, \text{lavg}}(t) = \frac{1}{|\mathcal{N}_i|} \sum_{j \in \mathcal{N}_i} \hat{y}_j(t)$ is the local public opinion specific to individual i , and includes only the expressed opinions of i 's neighbours. The cardinality of i 's set of neighbours \mathcal{N}_i is given by $|\mathcal{N}_i|$. From here on, \hat{y}_{avg} is referred to as the global public

opinion if there is a need to differentiate from the local public opinion $\hat{y}_{i,\text{avg}}$. One of the key results of Chapter 4 is to show that significant differences in the equilibrium opinion distributions of the network, in the context of pluralistic ignorance, can arise depending on whether Eq. (3.2) or Eq. (3.4) is used to model the expressed opinion.

Note that it is possible to further generalise Eq. (3.4) so that the local public opinion is $\hat{y}_{i,\text{avg}}(t) = \sum_{j \in \mathcal{N}_i} a_{ij} \hat{y}_j(t)$, where a_{ij} now become general weights satisfying $A \sim W$ and $\sum_{j=1}^n a_{ij} = 1, \forall i$. One potential, and reasonable, alternative to $a_{ij} = |\mathcal{N}_i|^{-1}$ is to set $a_{ij} = w_{ij}, \forall i, j \in \mathcal{I}$. The theoretical results detailed in Section 3.3 are not changed, and so no further exploration of the choice of weights of the local public opinion is taken. Rather, focus is placed on the differences in behaviour that may arise between the global and local public opinion models.

3.2.3 Obtaining a Compact Form for the Influence Network

To better analyse the system, a compact form for the opinion dynamics on the influence network $\mathcal{G}[W]$ is now obtained. Let $\mathbf{y} = [y_1, y_2, \dots, y_n]^\top$ and $\hat{\mathbf{y}} = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n]^\top$ be the stacked vectors of private and expressed opinions y_i and \hat{y}_i of the n individuals in the influence network $\mathcal{G}[W]$, respectively. For convenience, define $\boldsymbol{\phi} = [\phi_1, \dots, \phi_n]^\top$ and $\boldsymbol{\lambda} = [\lambda_1, \dots, \lambda_n]^\top$. The influence matrix W can be decomposed as $W = \widetilde{W} + \widehat{W}$ where $\widetilde{W} = \text{diag}(w_{ii})$, i.e. a diagonal matrix with diagonal entries w_{ii} . The matrix \widehat{W} is then a matrix with zero diagonal and with the same offdiagonal entries as W . That is, $\widehat{w}_{ij} = w_{ij}$ for all $j \neq i$ and $\widehat{w}_{ii} = 0$ for all i . Define $\Lambda = \text{diag}(\lambda_i)$ and $\Phi = \text{diag}(\phi_i)$. Using Eq. (3.2) and Eq. (3.3), one can verify that the compact form of the opinion dynamical system is given by

$$\begin{bmatrix} \mathbf{y}(t+1) \\ \hat{\mathbf{y}}(t) \end{bmatrix} = \begin{bmatrix} \Lambda (\widetilde{W} + \widehat{W}\Phi) & \Lambda \widehat{W} (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \\ \Phi & (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \end{bmatrix} \begin{bmatrix} \mathbf{y}(t) \\ \hat{\mathbf{y}}(t-1) \end{bmatrix} + \begin{bmatrix} (I_n - \Lambda) \mathbf{y}(0) \\ \mathbf{0}_n \end{bmatrix}. \quad (3.5)$$

As stated earlier, the initialisation is $\hat{\mathbf{y}}(0) = \mathbf{y}(0)$. Thus, for all $i \in \mathcal{I}$, there holds

$$y_i(1) = \lambda_i w_{ii} y_i(0) + \lambda_i \sum_{j \neq i}^n w_{ij} y_j(0) + (1 - \lambda_i) y_i(0), \quad (3.6)$$

and Eq. (3.5) holds for $t \geq 1$. The appearance of different time arguments within the one vector is dealt with by defining $\mathbf{x}_1(k) = \mathbf{y}(t)$ and $\mathbf{x}_2(k) = \hat{\mathbf{y}}(t-1)$. It follows that

$$\begin{bmatrix} \mathbf{x}_1(k+1) \\ \mathbf{x}_2(k+1) \end{bmatrix} = \begin{bmatrix} \Lambda (\widetilde{W} + \widehat{W}\Phi) & \Lambda \widehat{W} (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \\ \Phi & (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(k) \\ \mathbf{x}_2(k) \end{bmatrix} + \begin{bmatrix} (I_n - \Lambda) \mathbf{x}_1(0) \\ \mathbf{0}_n \end{bmatrix}. \quad (3.7)$$

The appropriate initialisation method is then to set $\mathbf{x}_2(1) = \mathbf{x}_1(0)$, giving $\mathbf{x}_1(1) = \Lambda W \mathbf{x}_1(0) + (I_n - \Lambda) \mathbf{x}_1(0)$, and study Eq. (3.7) for $k = 1, 2, \dots$. Last, define for future

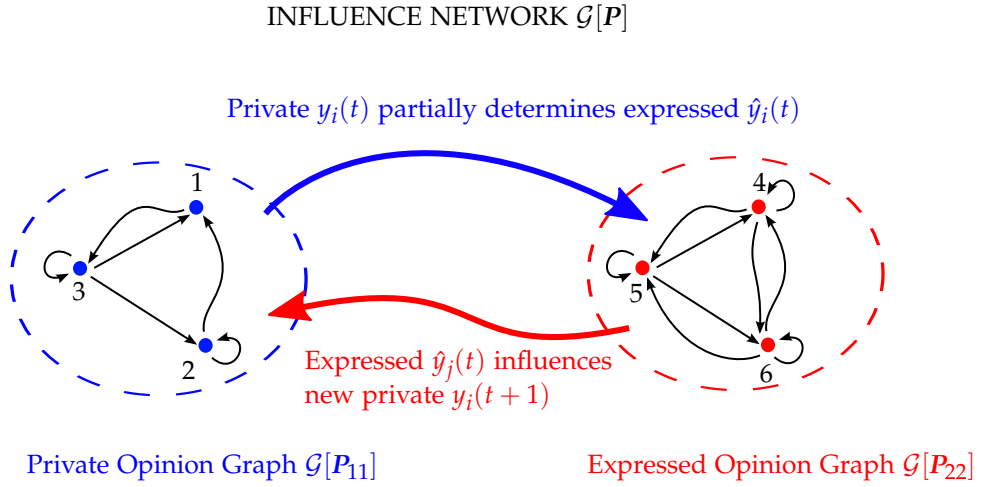


Figure 3.3: The influence network $\mathcal{G}[W]$ of n individuals, with each individual's private and expressed opinions evolving according to Eq. (3.1) and Eq. (3.2), is equivalent to the influence network $\mathcal{G}[P]$ with $2n$ nodes, evolving according to the system dynamics Eq. (3.5). Nodes $\mathcal{V}_p = \{v_1, \dots, v_n\}$ are associated with the private opinions y , and induce the *private opinion subgraph* $\mathcal{G}[P_{11}]$. Nodes $\mathcal{V}_e = \{v_{n+1}, \dots, v_{2n}\}$ are associated with the expressed opinions \hat{y} , and induce the *expressed opinion subgraph* $\mathcal{G}[P_{22}]$. The off-diagonal terms P_{12} and P_{21} describe influences (edges) between nodes in \mathcal{V}_p and \mathcal{V}_e . Here, an illustrative example with $n = 3$ is shown.

reference

$$P = \begin{bmatrix} \Lambda (\widetilde{W} + \widehat{W}\Phi) & \Lambda \widehat{W} (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \\ \Phi & (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}. \quad (3.8)$$

The influence network $\mathcal{G}[W]$ with each individual's private and expressed opinions evolving according to Eq. (3.1) and Eq. (3.2) can be viewed as a larger influence network $\mathcal{G}[P]$ with $2n$ nodes. This interpretation is explored in Fig. 3.3.

Following steps similar to those above, one can show that the evolution of opinions over the influence network, with Eq. (3.4) replacing Eq. (3.2) for every individual, is given by

$$\begin{bmatrix} y(t+1) \\ \hat{y}(t) \end{bmatrix} = \begin{bmatrix} \Lambda (\widetilde{W} + \widehat{W}\Phi) & \Lambda \widehat{W} (I_n - \Phi) A \\ \Phi & (I_n - \Phi) A \end{bmatrix} \begin{bmatrix} y(t) \\ \hat{y}(t-1) \end{bmatrix} + \begin{bmatrix} (I_n - \Lambda) y(0) \\ \mathbf{0}_n \end{bmatrix}, \quad (3.9)$$

where the i^{th} row of A has entries such that $w_{ij} = 0 \Leftrightarrow a_{ij} = 0$ and $w_{ij} > 0 \Leftrightarrow a_{ij} = \frac{1}{|\mathcal{N}_i|}$. In other words, $A \sim W$ (see Section 2.1). Thus, $\mathcal{G}[A]$ is a strongly connected and aperiodic graph where, for any given node i , the incoming edges have equal weight $a_{ij} = |\mathcal{N}_i|^{-1}$. It follows that A is row-stochastic, and because $\mathcal{G}[A]$ is strongly

connected and aperiodic, that A is primitive. For future purposes, define

$$\mathbf{Z} = \begin{bmatrix} \Lambda (\widehat{\mathbf{W}} + \widehat{\mathbf{W}}\Phi) & \Lambda \widehat{\mathbf{W}} (\mathbf{I}_n - \Phi) \mathbf{A} \\ \Phi & (\mathbf{I}_n - \Phi) \mathbf{A} \end{bmatrix} \quad (3.10)$$

Having obtained a compact form for the system dynamics, attention will now turn to study of the temporal evolution of the opinion vectors $\mathbf{y}(t)$ and $\hat{\mathbf{y}}(t)$.

3.3 Convergence Properties of the Model

This section will establish a number of results, providing sufficient conditions for the opinions to converge to an equilibrium (as opposed to the system entering an oscillating trajectory or becoming unstable). More importantly, detailed analysis of the opinions at equilibrium will provide insightful conclusions on the role of resilience and stubbornness in determining the final expressed and private opinions. To aid the analysis of the model and establish the convergence properties of the opinion dynamical system, an assumption is now formally introduced.

Assumption 3.1. *The network $\mathcal{G}[\mathbf{W}]$ is strongly connected and aperiodic, and the influence matrix \mathbf{W} is row-stochastic. Furthermore, there holds $\lambda_i, \phi_i \in (0, 1)$, for all $i \in \mathcal{I}$.*

It should be noted that for the purpose of convergence analysis, this assumption can almost certainly be relaxed to include graphs which are not strongly connected, and for values of ϕ_i, λ_i in the closed interval $[0, 1]$. Relaxation of Assumption 3.1 would be interesting future work, but the focus in this chapter is to gain deeper insight into the social processes governing the opinion evolution, for which Assumption 3.1 is suitable. Two lemmas are introduced to establish several properties of \mathbf{P} and $(\mathbf{I}_{2n} - \mathbf{P})^{-1}$, which will later be used to help prove several results in this section. A third lemma is provided to establish an invariance property. The proofs of these lemmas are given in Section 3.5.

Lemma 3.1. *Suppose that Assumption 3.1 holds. Then, the matrix \mathbf{P} given in Eq. (3.8) is nonnegative, irreducible, and there holds $\rho(\mathbf{P}) < 1$.*

Lemma 3.2. *Suppose that Assumption 3.1 holds. Define $\mathbf{Q} = \mathbf{I}_{2n} - \mathbf{P}$, where \mathbf{P} is given in Eq. (3.8), with \mathbf{Q} decomposed as*

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n - \mathbf{P}_{11} & -\mathbf{P}_{12} \\ -\mathbf{P}_{21} & \mathbf{I}_n - \mathbf{P}_{22} \end{bmatrix}.$$

Then, \mathbf{Q}_{11} and \mathbf{Q}_{22} are nonsingular, and \mathbf{Q}^{-1} is a positive matrix which has the form

$$\mathbf{Q}^{-1} = \begin{bmatrix} (\mathbf{Q}_{11} - \mathbf{Q}_{12}\mathbf{Q}_{22}^{-1}\mathbf{Q}_{21})^{-1} & -\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12}(\mathbf{Q}_{22} - \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12})^{-1} \\ -\mathbf{Q}_{22}^{-1}\mathbf{Q}_{21}(\mathbf{Q}_{11} - \mathbf{Q}_{12}\mathbf{Q}_{22}^{-1}\mathbf{Q}_{21})^{-1} & (\mathbf{Q}_{22} - \mathbf{Q}_{21}\mathbf{Q}_{11}^{-1}\mathbf{Q}_{12})^{-1} \end{bmatrix}. \quad (3.11)$$

Moreover, $\mathbf{R} = (\mathbf{Q}_{11} - \mathbf{Q}_{12}\mathbf{Q}_{22}^{-1}\mathbf{Q}_{21})^{-1}(\mathbf{I}_n - \Lambda)$ and $\mathbf{S} = -\mathbf{Q}_{22}^{-1}\mathbf{Q}_{21}$ are invertible, positive row-stochastic matrices.

Next, an invariant set result is established. In particular, it is shown that given a set of initial conditions $\mathbf{y}(0)$, the set

$$\mathcal{S} = \{y_i, \hat{y}_i : \min_{k \in \mathcal{I}} y_k(0) \leq y_i, \hat{y}_i \leq \max_{j \in \mathcal{I}} y_j(0), i \in \mathcal{I}\} \quad (3.12)$$

is a positive invariant set of the system Eq. (3.5). This is a desirable property for any opinion dynamics model. Section 2.3.3 discussed how opinion y_i may be scaled so that $a, b \in \mathbb{R}$ represent the two extremes of the opinion spectrum. Then, if $y_i(0) \in [a, b]$ and the model has the below property, $y_i(t) \in [a, b]$ for all $t \geq 0$, and so the opinions are always well defined (it does not make sense for $y_i < a$ or $y_i > b$ if a, b are the extremes).

Lemma 3.3 (Invariant Set). *Consider a network $\mathcal{G}[\mathbf{W}]$ where each individual i 's opinions $y_i(t)$ and $\hat{y}_i(t)$ update according to Eq. (3.1) and Eq. (3.2), respectively. Suppose Assumption 3.1 holds, and that $\mathbf{y}(0) = \hat{\mathbf{y}}(0)$. Then, for all $t \geq 0$, there holds*

$$\max \left\{ \max_{i \in \mathcal{I}} y_i(t), \max_{j \in \mathcal{I}} \hat{y}_j(t) \right\} \leq \max_{i \in \mathcal{I}} y_i(0) = \max_{i \in \mathcal{I}} \hat{y}_i(0), \quad (3.13)$$

$$\min \left\{ \min_{i \in \mathcal{I}} y_i(t), \min_{j \in \mathcal{I}} \hat{y}_j(t) \right\} \geq \min_{i \in \mathcal{I}} y_i(0) = \min_{i \in \mathcal{I}} \hat{y}_i(0). \quad (3.14)$$

Note that the Friedkin–Johnsen model has the same property. In Section 3.5, a simple simulation counterexample is given to show that *there need not hold*

$$\max \left\{ \max_{i \in \mathcal{I}} y_i(t), \max_{j \in \mathcal{I}} \hat{y}_j(t) \right\} \geq \max \left\{ \max_{i \in \mathcal{I}} y_i(t+1), \max_{j \in \mathcal{I}} \hat{y}_j(t+1) \right\} \quad (3.15)$$

$$\min \left\{ \min_{i \in \mathcal{I}} y_i(t), \min_{j \in \mathcal{I}} \hat{y}_j(t) \right\} \leq \min \left\{ \min_{i \in \mathcal{I}} y_i(t+1), \min_{j \in \mathcal{I}} \hat{y}_j(t+1) \right\} \quad (3.16)$$

for all $t \geq 0$. This is a semi-contractive property held by the DeGroot model, see Lemma 2.1. The main stability theorem, and a subsequent corollary for a condition to guarantee a consensus of opinions, are now presented.

Theorem 3.1 (Exponential Stability). *Consider a network $\mathcal{G}[\mathbf{W}]$ where each individual i 's opinions $y_i(t)$ and $\hat{y}_i(t)$ update according to Eq. (3.1) and Eq. (3.2), respectively. Suppose Assumption 3.1 holds, and that $\mathbf{y}(0) = \hat{\mathbf{y}}(0)$. Then, the system Eq. (3.5) converges exponentially fast to a unique equilibrium, which is given as*

$$\lim_{t \rightarrow \infty} \mathbf{y}(t) \triangleq \mathbf{y}^* = \mathbf{R}\mathbf{y}(0) \quad (3.17)$$

$$\lim_{t \rightarrow \infty} \hat{\mathbf{y}}(t) \triangleq \hat{\mathbf{y}}^* = \mathbf{S}\mathbf{y}^*, \quad (3.18)$$

where $\mathbf{R}, \mathbf{S} \in \mathbb{R}^{n \times n}$ are positive, row-stochastic matrices defined in Lemma 3.2.

Proof. Lemma 3.1 established that the time-invariant matrix \mathbf{P} satisfies $\rho(\mathbf{P}) < 1$. Standard linear systems theory [Rugh, 1996] is used to conclude that the linear, time-invariant system Eq. (3.5), with constant input $[((\mathbf{I}_n - \mathbf{A})\mathbf{y}(0))^\top, \mathbf{0}_n^\top]^\top$, converges

exponentially fast to the following unique equilibrium

$$\begin{bmatrix} \lim_{t \rightarrow \infty} \mathbf{y}(t) \\ \lim_{t \rightarrow \infty} \hat{\mathbf{y}}(t) \end{bmatrix} \triangleq \begin{bmatrix} \mathbf{y}^* \\ \hat{\mathbf{y}}^* \end{bmatrix} = (\mathbf{I}_{2n} - \mathbf{P})^{-1} \begin{bmatrix} (\mathbf{I}_n - \mathbf{\Lambda})\mathbf{y}(0) \\ \mathbf{0}_n \end{bmatrix} = \mathbf{Q}^{-1} \begin{bmatrix} (\mathbf{I}_n - \mathbf{\Lambda})\mathbf{y}(0) \\ \mathbf{0}_n \end{bmatrix}. \quad (3.19)$$

Having calculated the form of \mathbf{Q}^{-1} in Eq. (3.11), it can be verified that $\mathbf{y}^* = \mathbf{R}\mathbf{y}(0)$ and $\hat{\mathbf{y}}^* = \mathbf{S}\mathbf{R}\mathbf{y}(0) = \mathbf{S}\mathbf{y}^*$. Here, the definitions of \mathbf{R} and \mathbf{S} are given in Lemma 3.2, which also proved their positivity and row-stochasticity. This completes the proof. \square

This result indicates that the final private and expressed opinions depend on the initial private opinions, while the initial expressed $\hat{\mathbf{y}}(0)$ are forgotten exponentially fast. Thus, one could initialise $\hat{\mathbf{y}}(0)$ arbitrarily, and the final steady-state will be the same, though the transient will differ and the conclusions of Lemma 3.3 may no longer apply. The row-stochastic nature of \mathbf{R} and \mathbf{S} implies that the final private opinions (respectively final expressed opinions) are a convex combination of the initial private opinions (respectively the final private opinions). Additionally, positivity of \mathbf{R} and \mathbf{S} means every individual i 's initial $y_i(0)$ has an influence in the determination of any other individual j 's final opinions y_j^* and \hat{y}_j^* , a reflection of the strong connectedness of the network interactions. The following corollary establishes a condition for consensus of opinions, though it must be noted that part of the hypothesis for Theorem 3.1 is discarded.

Corollary 3.1 (Consensus of Opinions). *Suppose that $\phi_i \in (0,1)$, and $\lambda_i = 1$, for all $i \in \mathcal{I}$. Suppose further that $\mathcal{G}[\mathbf{W}]$ is strongly connected and aperiodic. Then, for the system Eq. (3.5), all opinions converge exponentially fast to a consensus, i.e. $\lim_{t \rightarrow \infty} \mathbf{y}(t) = \lim_{t \rightarrow \infty} \hat{\mathbf{y}}(t) = \alpha \mathbf{1}_n$ for some finite $\alpha \in \mathbb{R}$.*

Proof. Under the corollary assumptions, $\mathbf{\Lambda} = \mathbf{I}_n$, which implies that Eq. (3.7) becomes $\mathbf{x}(t+1) = \mathbf{P}\mathbf{x}(t)$. It was established in the proof of Lemma 3.1 (see Section 3.5.1) that \mathbf{P} has row sum equal to one, i.e. \mathbf{P} is nonnegative and row-stochastic. The lemma proof also established that $\mathcal{G}[\mathbf{P}]$ is strongly connected and aperiodic, and this remains unchanged when $\mathbf{\Lambda} = \mathbf{I}_n$. Theorem 2.1 may then be used conclude that consensus is achieved exponentially fast, i.e. $\lim_{t \rightarrow \infty} \mathbf{x}(t) = \alpha \mathbf{1}_{2n}$ for some $\alpha \in \mathbb{R}$. Recalling the definition of $\mathbf{x} = [\mathbf{x}_1^\top, \mathbf{x}_2^\top]^\top$ yields $\mathbf{y}(\infty) = \hat{\mathbf{y}}(\infty) = \alpha \mathbf{1}_n$. \square

3.3.1 Causes of Persistent Disagreement and Differences in Opinions

In this section, a result is obtained on the disagreement among the opinions at equilibrium. A key conclusion is that stubbornness and resilience create different expressed and private opinions in the same individual. This result and the social connotations are discussed, before the proof is presented.

Theorem 3.2 (Disagreement). *Suppose that the hypotheses in Theorem 3.1 hold. If there is no consensus of the initial opinions, i.e. $\mathbf{y}(0) \neq \alpha \mathbf{1}_n$ for some $\alpha \in \mathbb{R}$, then the final opinions*

obey the following inequalities

$$\max_{k \in \mathcal{I}} y_k(0) = \bar{y}_0 > \max_{i \in \mathcal{I}} y_i^* = y_{\max}^* > \max_{j \in \mathcal{I}} \hat{y}_j^* = \hat{y}_{\max}^* \quad (3.20a)$$

$$\min_{k \in \mathcal{I}} y_k(0) = \underline{y}_0 < \min_{i \in \mathcal{I}} y_i^* = y_{\min}^* < \min_{j \in \mathcal{I}} \hat{y}_j^* = \hat{y}_{\min}^* \quad (3.20b)$$

and $\hat{y}_{\min}^* \neq \hat{y}_{\max}^*$. Moreover, given a network $\mathcal{G}[\mathbf{W}]$ and parameter vectors $\boldsymbol{\phi}$ and $\boldsymbol{\lambda}$, the set of initial conditions $\mathbf{y}(0)$ for which $m > 0$ specific individuals $i_j \in \{i_1, \dots, i_m\} \subseteq \mathcal{I}$ have $y_{i_j}^* = \hat{y}_{i_j}^*$, i.e. $m \triangleq |\{i \in \mathcal{I} : y_i^* = \hat{y}_i^*\}|$, lies in a subspace of \mathbb{R}^n with dimension $n - m$.

This result shows that for generic initial conditions there is a persistent disagreement of opinions at the steady-state. This is a consequence of individuals not being maximally susceptible to influence, $\lambda_i < 1 \forall i \in \mathcal{I}$. If on the other hand $\lambda_i = 1$ for all $i \in \mathcal{I}$ then a consensus of opinions is reached exponentially fast (see Corollary 3.1). A second observation is that for any individual i in the network, $y_i^* \neq \hat{y}_i^*$ generically, which is a subtle but significant difference from the result stated in Eq. (3.20). That is, *the presence of both stubbornness and pressure to conform, and the strong connectedness of the network leads to an individual having different private and expressed opinions* in generic networks and with generic initial conditions¹. Without stubbornness ($\lambda_i = 1$ for all i), a consensus of opinions is reached, and without a pressure to conform ($\phi_i = 1$ for all i), an individual has the same private and expressed opinions. Without strong connectedness, some individuals will not be influenced to change opinions. One further consequence of Eq. (3.20) is that $y_{\max}^* - y_{\min}^* > \hat{y}_{\max}^* - \hat{y}_{\min}^*$, which implies that the level of agreement is greater among the expressed opinions when compared to the private opinions. In other words, individuals are more willing to agree with others when they are expressing their opinions in a social network due to a pressure to conform. Moreover, the extreme final expressed opinions are upper and lower bounded by the final private opinions, which are in turn upper and lower bounded by the extreme initial private opinions, clearly showing the effects of interpersonal influence and a pressure to conform.

The above conclusions continue to hold if only local public opinions are available, i.e. if Eq. (3.2) is replaced with Eq. (3.4), but key differences do exist (see Section 3.3.4).

Proof. If $\mathbf{y}(0) = \alpha \mathbf{1}_n$, for some finite $\alpha \in \mathbb{R}$ (i.e. the initial private opinions are at a consensus), then $\mathbf{y}^* = \hat{\mathbf{y}}^* = \alpha \mathbf{1}_n$ because \mathbf{R} and \mathbf{S} are row-stochastic. In what follows, it will be proved that if the initial private opinions are not at a consensus, then there is disagreement at equilibrium. It is suggested that the reader become familiar with the performance function $V(\mathbf{x})$ and coefficient of ergodicity $\tau(\mathbf{A})$ in Appendix A.1.1, as these will be used frequently in this proof.

First, the fact that $y_{\min}^* \neq y_{\max}^*$ is established. Note that $V(\mathbf{y}^*) = 0$ if and only if the private opinions are at a consensus, i.e. $\mathbf{y}^* = \beta \mathbf{1}_n$, for some $\beta \in \mathbb{R}$. Next, observe that $\mathbf{y}^* = \beta \mathbf{1}_n$ if and only if $\mathbf{R}\mathbf{y}(0) = \beta \mathbf{1}_n$, for some $\beta \in \mathbb{R}$. Note that \mathbf{R}

¹See Remark 3.3 below for further comments.

is invertible, because it is the product of two invertible matrices (see Lemma 3.2). Moreover, because R is row-stochastic, there holds $R\mathbf{1}_n = \mathbf{1}_n \Leftrightarrow R^{-1}R\mathbf{1}_n = R^{-1}\mathbf{1}_n \Leftrightarrow R^{-1}\mathbf{1}_n = \mathbf{1}_n$. That is, R^{-1} has row sum equal to one. Thus, premultiplying by R^{-1} on both sides of $R\mathbf{y}(0) = \beta\mathbf{1}_n$ yields $\mathbf{y}(0) = \beta R^{-1}\mathbf{1}_n = \beta\mathbf{1}_n$. In other words, the only possibility for there to be a consensus of the final private opinions, $\mathbf{y}^* = \beta\mathbf{1}_n$, is if the initial private opinions are at a consensus. Recalling the hypothesis of the theorem statement (that $\mathbf{y}(0)$ is not equal to $\alpha\mathbf{1}_n$, for some $\alpha \in \mathbb{R}$), it is immediately clear that \mathbf{y}^* is not at a consensus. Thus, $y_{\min}^* \neq y_{\max}^*$ as claimed.

Next, the inequalities Eq. (3.20a) and Eq. (3.20b) are proved. The fact that R, S are positive and row-stochastic implies that $\tau(R) < 1$ and $\tau(S) < 1$. Because R is invertible, R cannot be a matrix where all the rows are equal, i.e. $R \neq \mathbf{1}_n \mathbf{z}^\top$ for some $\mathbf{z} \in \mathbb{R}^n$. This means that $\tau(R) > 0$ (see below Eq. (A.9) in Appendix A). Similarly, one can prove that $\tau(S) > 0$. In the above paragraph, it was shown that if there is no consensus of the initial private opinions, then there is no consensus of the final initial opinions, i.e. $V(\mathbf{y}^* = R\mathbf{y}(0)) > 0$. Thus, by recalling that $V(A\mathbf{x}) \leq \tau(A)V(\mathbf{x})$ (see Eq. (A.11) in Appendix A) and the above facts, one can obtain the conclusion that $0 < V(\mathbf{y}^* = R\mathbf{y}(0)) < V(\mathbf{y}(0))$, which establishes the left hand inequality of both Eq. (3.20a) and Eq. (3.20b). Following steps similar to the above, but which are omitted, one can show that $0 < V(\hat{\mathbf{y}}^* = S\mathbf{y}^*) < V(\mathbf{y}^*)$, which establishes the right hand inequality of both Eq. (3.20a) and Eq. (3.20b), and also establishes the fact that $\hat{y}_{\min}^* \neq \hat{y}_{\max}^*$. This means there is disagreement in the final expressed opinions.

Last, it remains to prove² that for generic initial conditions, $y_i^* \neq \hat{y}_i^*$. First, the case for $m \geq 2$ will be proved. Consider the final private opinions. From $\mathbf{y}^* = R\mathbf{y}(0)$, it is clear that $y_i^* = y_j^*$ for any $i, j \in \mathcal{I}$ and $i \neq j$, if and only if $(\mathbf{r}_i - \mathbf{r}_j)^\top \mathbf{y}(0) = 0$, where \mathbf{r}_i^\top is the i^{th} row of R . Since R is invertible, $\text{rank}(R) = n$, and thus $\mathbf{r}_i \neq \mathbf{r}_j$ for any $i \neq j$. This implies that $\mathbf{y}(0)$ lies on a hyperplane of dimension $n - 1$ orthogonal to $\mathbf{r}_i - \mathbf{r}_j \neq \mathbf{0}_n$ (if $\mathbf{r}_i = \mathbf{r}_j$ then the space \mathbb{R}^n would be orthogonal to $\mathbf{y}(0)$). Recall that R is determined uniquely by W, ϕ and λ , and is independent of $\mathbf{y}(0)$. Thus, for a given network with a given set of parameters W, ϕ and λ , one is able to conclude that the set of initial conditions for which $m \geq 2$ individuals have the same private opinion lies in a subspace of \mathbb{R}^n of dimension $n - m$.

Let $\mathcal{H}_1, \dots, \mathcal{H}_H$ be the H disjoint sets of individuals $i \in \mathcal{I}$ who have the same final private opinion, with the private opinion value for each disjoint set being different. That is, for $h \in \{1, \dots, H\}$, $\mathcal{H}_h = \{i_h, j_h \in \mathcal{I} : y_{i_h}^* = y_{j_h}^*\}$ with $y_{i_p}^* \neq y_{i_q}^*$ for all $p, q \in \{1, \dots, H\}$ and $p \neq q$. Denote the cardinality of \mathcal{H}_h by $m_h = |\mathcal{H}_h|$, and denote the set $\mathcal{K} = \{i \in \mathcal{I} : y_i^* = \hat{y}_i^*\}$, with cardinality $K = |\mathcal{K}|$. From Eq. (3.2), it follows trivially that for any $i \in \mathcal{I}$, $y_i^* = \hat{y}_i^*$ if and only if $y_i^* = \hat{y}_{\text{avg}}^*$. It then follows that

²During the thesis examination process, one examiner identified a different proof to show that generically, $y_i^* \neq \hat{y}_i^*$, the summary of which is provided here. First, one observes that $\hat{y}_i^* = y_i^* \Rightarrow \hat{y}_{\text{avg}}^* = \mathbf{1}_n^\top \hat{\mathbf{y}}^* / n$. Thus, $\hat{y}_i^* = y_i^*$ for m specific individuals if and only if $\hat{\mathbf{y}}^*$ lies in an $n - m$ -dimensional subspace of \mathbb{R}^n , denoted as \mathcal{D} , because there must be m independent equations satisfying $(\mathbf{e}_i - \frac{1}{n}\mathbf{1}_n)^\top \mathbf{y}^* = 0$. From Lemma 3.2, one has $\mathbf{y}^* = R\mathbf{S}\mathbf{y}(0)$. It follows that $\hat{y}_i^* = y_i^*$ for m specific individuals only if $\mathbf{y}(0)$ belongs to the inverse image (by RS) of \mathcal{D} ; the inverse image has dimension $n - m$ because R, S are invertible.

$K \leq \max_{h \in \{1, \dots, H\}} m_h$. Thus, for a given network with a given set of parameters \mathbf{W} , $\boldsymbol{\phi}$ and $\boldsymbol{\lambda}$, the set of initial conditions $\mathbf{y}(0)$ for which $m \geq 2$ individuals i have $y_i^* = \hat{y}_i^*$ lies in a subspace of \mathbb{R}^n with dimension of at most $n - m$.

The case of $m = 1$ will now be proved, i.e. there is a single individual i in the network with $y_i^* = \hat{y}_i^* = \hat{y}_{\text{avg}}^*$. One has $\hat{y}_{\text{avg}}^* = \tilde{\mathbf{r}}^\top \mathbf{y}(0)$, where $\tilde{\mathbf{r}}^\top = \frac{1}{n} \mathbf{1}_n^\top \mathbf{S} \mathbf{R}$, which implies that $y_i^* = \hat{y}_i^*$ if and only if $(\mathbf{r}_i - \tilde{\mathbf{r}})^\top \mathbf{y}(0) = 0$. In other words, $\mathbf{r}_i - \tilde{\mathbf{r}}$ must be orthogonal to $\mathbf{y}(0)$. First, it must be proved that $\mathbf{r}_i \neq \tilde{\mathbf{r}}$. Define $\tilde{\mathbf{s}}^\top = \frac{1}{n} \mathbf{1}_n^\top \mathbf{S}$, and observe that $\tilde{\mathbf{s}}^\top \mathbf{R} = \tilde{\mathbf{r}}^\top$ is a linear combination of the rows of \mathbf{R} , denoted as $\mathbf{r}_1^\top, \dots, \mathbf{r}_n^\top$. Thus, $\mathbf{r}_i^\top = \tilde{\mathbf{s}}^\top \mathbf{R}$ if and only if \mathbf{r}_i^\top is a linear combination of the rows of \mathbf{R} . But since \mathbf{R} is invertible, then clearly \mathbf{r}_i^\top cannot be a linear combination of $\mathbf{r}_1^\top, \dots, \mathbf{r}_n^\top$ (otherwise \mathbf{R} would not have full rank). Since \mathbf{S} and \mathbf{R} are independent of $\mathbf{y}(0)$ then so is $\tilde{\mathbf{r}}$, and it follows that for a network with a given set of parameters \mathbf{W} , $\boldsymbol{\phi}$, $\boldsymbol{\lambda}$, the set of $\mathbf{y}(0)$ which leads to a single individual i having $y_i^* = \hat{y}_i^*$ is in a subspace of \mathbb{R}^n of dimension $n - 1$. This completes the proof. \square

Remark 3.3. From Theorem 3.2 and its proof, one can then say that for generic networks and generic initial conditions, there will be no individuals who have the same final private opinions, and no individual will have the same final private and expressed opinion. Suppose that the parameters defining the network are given, i.e. \mathbf{W} , $\boldsymbol{\phi}$ and $\boldsymbol{\lambda}$. Suppose further that each $y_i(0), i \in \mathcal{I}$ is sampled from some continuous distribution (which might be a uniform, normal, beta, etc.). If the distribution is over a non-degenerate interval³ (which may be infinite), one can then see that, with probability zero, there will be $m > 0$ individuals that have $y_i^* = \hat{y}_i^*$. In other words, if one runs p experiments in which one independently selects initial condition vectors $\mathbf{y}(0)$ from a non-degenerate distribution, and if q is the number of those experiments which result in $y_i^* = \hat{y}_i^*$ for some $i \in \mathcal{I}$, then $\lim_{p \rightarrow \infty} q/p = 0$. This can be viewed from another perspective: the set of $\mathbf{y}(0)$ for which $y_i^* = \hat{y}_i^*$ for some $i \in \mathcal{I}$ belongs in a subspace of \mathbb{R}^n of at most dimension $n - 1$, and any such subspace has a Lebesgue measure of zero. The same can be said for the initial conditions which lead to $y_i^* = y_j^*$ for $i \neq j$.

3.3.2 Estimating Disagreement in the Private Opinions

This section presents a method for estimating the disagreement among the private opinions at equilibrium given limited knowledge of the network. The social implications of the result are discussed, with the proof presented in Appendix Section 3.5

Corollary 3.2 (Estimating Private Disagreement). *Suppose that the hypotheses in Theorem 3.1 hold. Then, there holds*

$$\frac{\hat{y}_{\max}^* - \hat{y}_{\min}^*}{\kappa(\boldsymbol{\phi})} \leq y_{\max}^* - y_{\min}^* \quad (3.21)$$

where $\kappa(\boldsymbol{\phi}) = 1 - \frac{\phi_{\min}}{\phi_{\max}}(1 - \phi_{\max}) \in (0, 1)$ and $\phi_{\max} = \max_{i \in \mathcal{I}} \phi_i$, $\phi_{\min} = \min_{i \in \mathcal{I}} \phi_i$.

³A statistical distribution is degenerate if the cumulative distribution function $F(x, k_0) = 0$ if $x < k_0$ and $F(x, k_0) = 1$ if $x \geq k_0$.

For the purposes of monitoring the level of unvoiced discontent in a network (e.g. to prevent drastic and unforeseen actions or violence [Goodwin, 2011; Kuran, 1989; Duggins, 2017]), it is of interest to obtain more knowledge about the level of disagreement among the private opinions of the individuals in the network. This can be represented by $y_{\max}^* - y_{\min}^*$. A fundamental issue is that such information is unlikely to be obtainable (by definition private opinions are not readily observable except in certain situations like the post-experimental interviews conducted by Asch in his experiments, see Chapter 4). On the other hand, for a given influence network, one expects that the level of disagreement $\hat{y}_{\max}^* - \hat{y}_{\min}^*$ is readily available (by definition opinions are expressed). While one cannot expect to know every ϕ_i , the author argues that ϕ_{\max} and ϕ_{\min} might be obtained, if not accurately then approximately. Corollary 3.2 therefore gives a method for computing a *lower bound* on the level of private disagreement given limited knowledge of (i) the final expressed opinions, and (ii) an estimate of the resilience levels of the individuals.

It is obvious that if $\kappa(\phi)$ is small, then even strong agreement among the expressed opinions (small $\hat{y}_{\max}^* - \hat{y}_{\min}^*$) does not preclude significant disagreement in the final private opinions of the individuals. A small $\kappa(\phi)$ may occur if ϕ_{\max} is small and the ratio ϕ_{\min}/ϕ_{\max} is close to 1. One example is a highly structured group with rigid protocols, such as an authoritarian government.

Proof. See Section 3.5.5. □

Remark 3.4 (Tightness of the bound). *Key to Corollary 3.2 is the proof that the coefficient of ergodicity for S is bounded from above as $\tau(S) \leq \kappa(\phi)$. The tightness of this bound depends on the ratio ϕ_{\min}/ϕ_{\max} ; the closer the ratio is to one (i.e. as the “force” of the pressure to conform felt by each individual becomes more uniform), the tighter the bound. If $\phi_{\min}/\phi_{\max} = 1$, i.e. all resilience values are equal, then $\tau(S) = \kappa(\phi)$. This can be concluded by examining the proof, and noting that the key inequalities in Eq. (3.37) and Eq. (3.38) involve ϕ_{\min} and ϕ_{\max} . If $\phi_i \forall i$ are known, one can obtain $y^* = S^{-1}\hat{y}^*$ precisely.*

3.3.3 An Individual’s Resilience Affects Everyone

An interesting result is now presented, that shows how individual i ’s resilience to the pressure to conform, as measured by ϕ_i , has a effect on the final expressed opinions of all other individuals. In particular, i ’s resilience is propagated through the public opinion and the strongly connected network. Again, a discussion of the theoretical result from a sociological context is given immediately after the corollary statement, with the proof provided last.

Corollary 3.3 (Individual Resilience). *Suppose that the hypotheses in Theorem 3.1 hold. Then, the matrix $S \in \mathbb{R}^{n \times n}$ appearing in Eq. (3.18) is a function of $\phi_i, i \in \mathcal{I}$ and has partial*

derivative $\frac{\partial(S)}{\partial\phi_i} \in \mathbb{R}^{n \times n}$ with the following sign pattern

$$\frac{\partial(S)}{\partial\phi_i} = \begin{matrix} & \begin{matrix} i^{th} column \end{matrix} \\ \begin{matrix} \frac{\partial(S)}{\partial\phi_i} \end{matrix} = \begin{bmatrix} - & - & \dots & - & + & - & \dots & - \\ - & - & \dots & - & + & - & \dots & - \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & - \\ - & - & \dots & - & + & - & \dots & - \\ - & - & \dots & - & + & - & \dots & - \end{bmatrix} \end{matrix}. \quad (3.22)$$

That is, $\frac{\partial(S)}{\partial\phi_i}$ has positive entries in the i^{th} column and all other entries are negative.

From the fact that $\hat{y}^* = S y^*$, where S is positive and row-stochastic, it follows that individual k 's final expressed opinion \hat{y}_k^* is a convex combination of all individuals' final private opinions y_j^* , with convex weights s_{kj} , $j = 1, \dots, n$. Intuitively, increasing ϕ_k makes individual k more resilient to the pressure to conform, and this is confirmed by the above; $\frac{\partial s_{kk}}{\partial\phi_k} > 0$ and $\frac{\partial s_{kj}}{\partial\phi_k} < 0$ for any $j \neq k$ and thus $\hat{y}_k^* \rightarrow y_k^*$.

More importantly, the above result yields a surprising and nontrivial fact; *every entry* of the k^{th} column of $\frac{\partial(S)}{\partial\phi_k}$ is strictly positive, and all other entries of $\frac{\partial(S)}{\partial\phi_k}$ are strictly negative. In context, any change in individual k 's resilience directly impacts every other individual's final expressed opinion due to the network of interpersonal influences. In particular, as ϕ_k increases (decreases), an individual j 's final expressed opinion \hat{y}_j^* becomes more influenced by (less influenced by) the final private opinion y_k^* of individual k due to the convex weight s_{jk} increasing, since $\frac{\partial s_{jk}}{\partial\phi_k} > 0$ (decreasing, since $\frac{\partial s_{jk}}{\partial\phi_k} < 0$). Thus, one concludes that an individual (agent) level process of being resilient to conforming to a group norm can have wide reaching consequences on network-level dynamics. This will be highlighted in Chapter 4, where it will be shown how a few stubborn extremists with high resilience can create massively different expressed and private opinions among the general network population. Here, a mathematical treatment is provided to help explain the phenomena that is reported in Chapter 4.

Proof. First, one may verify that S is dependent only on ϕ , and S is invertible, and continuously differentiable, for all $\phi_i \in (0, 1)$. It follows from Lemma A.3 that

$$\frac{\partial S(\phi)}{\partial\phi_i} = -S(\phi) \left(\frac{\partial S^{-1}(\phi)}{\partial\phi_i} \right) S(\phi), \quad (3.23)$$

and note that in future, the argument ϕ will be dropped from $S(\phi)$ and $S^{-1}(\phi)$ when there is no confusion. First, note that $\frac{\partial \Phi^{-1}}{\partial\phi_i} = -\phi_i^{-2} \mathbf{e}_i \mathbf{e}_i^\top$, where \mathbf{e}_i is the i^{th} canonical unit vector (see Section 2.1). It can be proved that $S^{-1} = \Phi^{-1} - \Phi^{-1}(\mathbf{I}_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n}$ (see

Eq. (3.31) in Section 3.5.2), and by using Lemma A.3, one obtains

$$\begin{aligned}\frac{\partial S(\boldsymbol{\phi})}{\partial \phi_i} &= -S \left(-\frac{1}{\phi_i^2} \mathbf{e}_i \mathbf{e}_i^\top + \frac{1}{\phi_i^2} \mathbf{e}_i \mathbf{e}_i^\top \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right) S \\ &= \frac{1}{\phi_i^2} S \mathbf{e}_i \left(\mathbf{e}_i^\top - \frac{1}{n} \mathbf{1}_n^\top \right) S,\end{aligned}$$

with the second equality being obtained by noting that $\mathbf{e}_i^\top \mathbf{1}_n = 1$. It suffices to prove the corollary claim, if it can be shown that the row vector $(\mathbf{e}_i^\top - \frac{1}{n} \mathbf{1}_n^\top) S$ has a strictly positive i^{th} entry and all other entries are strictly negative. This is because $S > 0$ implies that the row vector $S \mathbf{e}_i > 0$. This will be achieved by showing that

$$\left(\mathbf{e}_i^\top - \frac{1}{n} \mathbf{1}_n^\top \right) S \mathbf{e}_i > 0 \quad (3.24)$$

$$\left(\mathbf{e}_i^\top - \frac{1}{n} \mathbf{1}_n^\top \right) S \mathbf{e}_j < 0, \quad \forall j \neq i. \quad (3.25)$$

Toward this end, the following useful quantity is first calculated:

$$\mathbf{e}_i^\top S^{-1} = \mathbf{e}_i^\top \left(\boldsymbol{\Phi}^{-1} - \boldsymbol{\Phi}^{-1} (I_n - \boldsymbol{\Phi}) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right) = \frac{1}{\phi_i} \mathbf{e}_i^\top - \frac{1}{n} \left(\frac{1}{\phi_i} - 1 \right) \mathbf{1}_n^\top,$$

which holds because $\mathbf{e}_i^\top \boldsymbol{\Phi}^{-1} = \phi_i^{-1} \mathbf{e}_i^\top$. Postmultiplying by S on both sides of the equation $\mathbf{e}_i^\top S^{-1} = \frac{1}{\phi_i} \mathbf{e}_i^\top - \frac{1}{n} \left(\frac{1}{\phi_i} - 1 \right) \mathbf{1}_n^\top$ yields $\mathbf{e}_i^\top = \frac{1}{\phi_i} \mathbf{e}_i^\top S - \frac{1}{n} \left(\frac{1}{\phi_i} - 1 \right) \mathbf{1}_n^\top S$. Rearranging this yields the following two equalities

$$\mathbf{e}_i^\top S = \phi_i \mathbf{e}_i^\top + \frac{1}{n} (1 - \phi_i) \mathbf{1}_n^\top S \quad (3.26)$$

$$\mathbf{1}_n^\top S = \frac{n}{1 - \phi_i} \left(\mathbf{e}_i^\top S - \phi_i \mathbf{e}_i^\top \right). \quad (3.27)$$

First, Eq. (3.25) will be proved. By using the equality of Eq. (3.26) for substitution, observe that the left hand side of Eq. (3.25) can be evaluated as

$$\begin{aligned}\left(\mathbf{e}_i^\top S - \frac{1}{n} \mathbf{1}_n^\top S \right) \mathbf{e}_j &= \left(\phi_i \mathbf{e}_i^\top + \frac{1}{n} (1 - \phi_i) \mathbf{1}_n^\top S - \frac{1}{n} \mathbf{1}_n^\top S \right) \mathbf{e}_j \\ &= -\frac{\phi_i}{n} \mathbf{1}_n^\top S \mathbf{e}_j,\end{aligned} \quad (3.28)$$

by making use of the fact that $\mathbf{e}_i^\top \mathbf{e}_j = 0$ for any $j \neq i$. Note that the quantity $\mathbf{1}_n^\top S \mathbf{e}_j$ is the sum of all entries of the j^{th} column of S , and this quantity is strictly positive because $S > 0$. Thus, $-\phi_i \mathbf{1}_n^\top S \mathbf{e}_j / n < 0$, which proves Eq. (3.25). Next, Eq. (3.24) will be proved. Using the equality of Eq. (3.27) for substitution, observe that the left hand

side of Eq. (3.24) is

$$\begin{aligned} (\mathbf{e}_i^\top \mathbf{S} - \frac{1}{n} \mathbf{1}_n^\top \mathbf{S}) \mathbf{e}_i &= \mathbf{e}_i^\top \mathbf{S} \mathbf{e}_i - \frac{1}{1 - \phi_i} (\mathbf{e}_i^\top \mathbf{S} \mathbf{e}_i - \phi_i \mathbf{e}_i^\top \mathbf{e}_i) \\ &= \frac{\phi_i}{1 - \phi_i} (1 - \mathbf{e}_i^\top \mathbf{S} \mathbf{e}_i) > 0. \end{aligned} \quad (3.29)$$

The inequality is obtained by observing that 1) $0 < \phi_i < 1 \Rightarrow \phi_i / (1 - \phi_i) > 0$, and 2) $1 - \mathbf{e}_i^\top \mathbf{S} \mathbf{e}_i > 0$ because $\mathbf{e}_i^\top \mathbf{S} \mathbf{e}_i$ is simply the i^{th} diagonal entry of the positive row-stochastic matrix \mathbf{S} . That is, $0 < \mathbf{e}_i^\top \mathbf{S} \mathbf{e}_i = s_{ii} < 1$. This proves Eq. (3.24), and the corollary proof is complete. \square

For completeness, note that $\frac{\partial \mathbf{S}(\phi)}{\partial \phi_i} \mathbf{1}_n = \mathbf{0}_n$, i.e. $\frac{\partial \mathbf{S}(\phi)}{\partial \phi_i}$ is a matrix with zero row sum. This property is related to the fact that \mathbf{S} is row-stochastic: any increase in an entry's magnitude must be matched with decreases in all other entries of the same row.

3.3.4 Local Public Opinions

As detailed in Section 3.2.2, one can consider a model where the global public opinion \hat{y}_{avg} is unavailable, but a local, individual specific, public opinion $\hat{y}_{i,\text{avg}}$ is available. In other words, Eq. (3.2) is replaced by Eq. (3.4). The previously detailed theoretical results all continue to hold with obvious adjustments to the statements, except for Corollary 3.2. No proofs are given, other than the following comments which point out the major differences in analysis with the global public opinion model.

Regarding Lemmas 3.1 and 3.2: Observe that to go from \mathbf{P} in Eq. (3.8) to \mathbf{Z} in Eq. (3.10), one simply replaces the row-stochastic and primitive matrix $\frac{\mathbf{1}_n \mathbf{1}_n^\top}{n}$ with the row-stochastic and primitive matrix \mathbf{A} (because \mathbf{W} is primitive under Assumption 3.1 and $\mathbf{A} \sim \mathbf{W}$, then \mathbf{A} is also primitive). This means that the result of Lemma 3.1 continues to hold if \mathbf{P} is replaced with \mathbf{Z} in the lemma statement, because all the properties of $\mathcal{G}[\frac{\mathbf{1}_n \mathbf{1}_n^\top}{n}]$ used in the proof (e.g. strong connectedness, and row-stochasticity of $\frac{\mathbf{1}_n \mathbf{1}_n^\top}{n}$) also hold for $\mathcal{G}[\mathbf{A}]$. It also means that Lemma 3.2 also applies for $\tilde{\mathbf{Q}} = \mathbf{I}_{2n} - \mathbf{Z}$. Again, the fact that \mathbf{A} is row-stochastic and primitive is all that is need to prove the lemma statement for $\tilde{\mathbf{Q}}$.

Regarding the remaining results: One can readily prove that Theorems 3.1 and 3.2, with the obvious adjustments, continue to hold, since the proofs primarily rely on Lemmas 3.1 and 3.2. Corollary 3.1 and Lemma 3.3 will also continue to hold. A key difference however, is that \mathbf{S} depends on \mathbf{A} , which in turn is dependent on \mathbf{W} (\mathbf{A} is dependent on the structure of \mathbf{W} , but not the value of its entries). This means that \mathbf{S} , i.e. the matrix whose entries form the convex weights which relate $\hat{\mathbf{y}}^*$ to \mathbf{y}^* , now depends on the structure of \mathbf{W} . Investigations of pluralistic ignorance in the following Chapter 4 will show that the effects of changes in \mathbf{A} can be substantial. Corollary 3.3 can also easily be extended. Note that $\mathbf{e}_i^\top \mathbf{A} = \mathbf{a}_i^\top$, where \mathbf{a}_i^\top is the i^{th} row of \mathbf{A} . Since \mathbf{A} is irreducible, then at least one entry of \mathbf{a}_i^\top is strictly positive. The proof method then follows identical steps, but with \mathbf{A} replacing $\frac{\mathbf{1}_n \mathbf{1}_n^\top}{n}$.

Note that Corollary 3.2 was proved using nontrivial calculations that made explicit use of the fact that the global public opinion updating Eq. (3.2) was captured by an all-to-all (or complete) influence network $\mathcal{G}[\frac{1_n 1_n^\top}{n}]$. With local public opinion updating Eq. (3.4), these calculations no longer hold. This suggests that obtaining a similar result, but for local public opinion updating, might be extremely difficult, but not necessarily impossible and may be a future direction of research.

3.3.5 Simulations

A simulation is now presented to illustrate the theoretical results of this chapter. A 3-regular network⁴ $\mathcal{G}[\mathbf{W}]$ with $n = 18$ is generated; the matrix \mathbf{W} is given in Section 3.5.6. Self-loops are added to each node (to ensure $\mathcal{G}[\mathbf{W}]$ is aperiodic), and the influence weights w_{ij} are obtained as follows. The value of each w_{ij} is drawn randomly from a uniform distribution in the interval $(0, 1)$ if $(v_j, v_i) \in \mathcal{E}$, and once all w_{ij} are determined, the weights are normalised by dividing all entries in row i by $\sum_{j=1}^n w_{ij}$. This ensures that \mathbf{W} is row-stochastic and nonnegative. For $i \neq j$, it is not required that $w_{ij} = w_{ji}$ (which would result in an undirected graph), but for simplicity and convenience the simulations impose⁵ that $w_{ij} > 0 \Leftrightarrow w_{ji} > 0$. The values of $y_i(0)$, ϕ_i , and λ_i , are picked from beta distributions, which have two parameters α and β . For $\alpha, \beta > 1$, a beta distribution of the variable x is unimodal and satisfies $x \in (0, 1)$, which is precisely what is required to satisfy Assumption 3.1 regarding ϕ_i, λ_i . The beta distribution parameters are (i) $\alpha = 2, \beta = 2$ for $y_i(0)$, (ii) $\alpha = 2, \beta = 2$ for ϕ_i , and (iii) $\alpha = 2, \beta = 8$ for λ_i .

The temporal evolution of opinions is shown in Fig. 3.4. Several of the results detailed in this chapter can be observed. In particular, it is clear that Eq. (3.18) holds. That is, there is no consensus of the expressed or private opinions at equilibrium. Moreover, the disagreement among the final expressed opinions, $\hat{y}_{\max}^* - \hat{y}_{\min}^*$, is strictly smaller than the disagreement among the final private opinions, $y_{\max}^* - y_{\min}^*$. Separate to this, the final private opinions enclose the final expressed opinions from above and below. For the given simulation, the largest and smallest resilience values are $\phi_{\max} = 0.9437$ and $\phi_{\min} = 0.1994$, respectively. This implies that $\kappa(\boldsymbol{\phi}) = 0.9881$. One can also obtain that $\hat{y}_{\max}^* - \hat{y}_{\min}^* = 0.1613$. From Eq. (3.21), this indicates that $y_{\max}^* - y_{\min}^* \geq 0.163$. The simulation result is consistent with the lower bound, in that $y_{\max}^* - y_{\min}^* = 0.3455$. Also, the bound is not tight, since ϕ_{\min}/ϕ_{\max} is far from 1 (see Remark 3.4).

For the same $\mathcal{G}[\mathbf{W}]$, with the same initial conditions $y_i(0)$ and resilience ϕ_i , a second simulation is run with $\lambda_1 = 1, \forall i \in \mathcal{I}$. As shown in Fig. 3.5, the opinions converge to a consensus $\mathbf{y}^* = \hat{\mathbf{y}}^* = \alpha \mathbf{1}_n$, for some $\alpha \in \mathbb{R}$, which illustrates Corollary 3.1.

⁴A k -regular graph is one which every node v_i has k neighbours, i.e. $|\mathcal{N}_i| = k \forall i \in \mathcal{I}$.

⁵Such an assumption is not needed for the theoretical results, but is a simple way to ensure that all directed graphs generated using the MATLAB package are strongly connected.

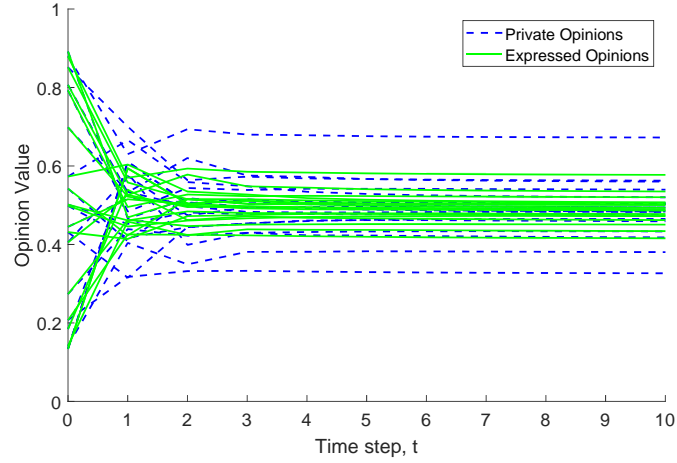


Figure 3.4: Temporal evolution of opinions for 18 individuals in an influence network. The green and dotted blue lines represent the expressed and private opinions of the individuals, respectively.

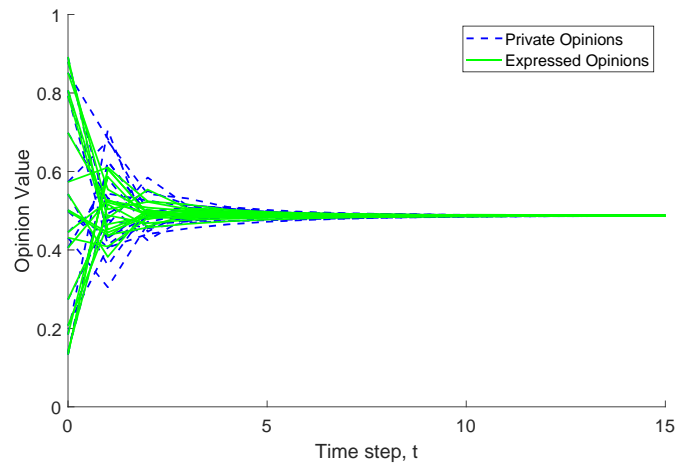


Figure 3.5: Temporal evolution of opinions for 18 individuals in an influence network. The green and dotted blue lines represent the expressed and private opinions of the individuals, respectively. The lack of stubbornness, $\lambda_i = 1, \forall i$, means that all opinions reach a consensus.

3.4 Conclusion

This chapter introduced an opinion dynamics model where each individual had a private and an expressed opinion on the same topic. In particular, the expressed opinion of an individual was altered from his/her private opinion due to a pressure to conform. A number of results were obtained, beginning with establishing of the dynamical properties of the opinions evolving on the influence network. Beyond the stability analysis, a number of semi-quantitative conclusions were drawn on the distributions of opinions at equilibrium, which in the social context of the model gave illuminating insight into how stubbornness and resilience affected the opinion evolution. A number of exciting possible future works are discussed in Chapter 10 after a further study of the model, focussing on the explanation of social psychological phenomena using the model, is conducted in Chapter 4.

3.5 Appendix: Proofs and Simulations

3.5.1 Proof of Lemma 3.1

First, it is easily verified that P is a nonnegative matrix by using the fact that W , Λ , $I_n - \Phi$, $\mathbf{1}_n \mathbf{1}_n^\top / n$ are all nonnegative matrices (since $0 < \phi_i, \lambda_i < 1$). Next, observe that

$$\begin{aligned} \begin{bmatrix} \Lambda(\widetilde{W} + \widehat{W}\Phi) & \Lambda\widehat{W}(I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \\ \Phi & (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \end{bmatrix} \begin{bmatrix} \mathbf{1}_n \\ \mathbf{1}_n \end{bmatrix} &= \begin{bmatrix} \Lambda(\widetilde{W} + \widehat{W}\Phi)\mathbf{1}_n + \Lambda\widehat{W}(I_n - \Phi)\mathbf{1}_n \\ \Phi\mathbf{1}_n + (I_n - \Phi)\mathbf{1}_n \end{bmatrix} \\ &= \begin{bmatrix} \Lambda(\widetilde{W} + \widehat{W})\mathbf{1}_n \\ \mathbf{1}_n \end{bmatrix} = \begin{bmatrix} \Lambda\mathbf{1}_n \\ \mathbf{1}_n \end{bmatrix}. \end{aligned} \quad (3.30)$$

with the first equality obtained by observing that $\frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \mathbf{1}_n = \mathbf{1}_n$, and with the last equality obtained by recalling that $W = \widetilde{W} + \widehat{W}$ is row-stochastic, i.e. $W\mathbf{1}_n = \mathbf{1}_n$. Because $\lambda_i < 1 \forall i$, it is immediately clear that Eq. (3.30) implies that rows $1, \dots, n$ of P each have row sum equal to a value strictly less than one, while rows $n+1, \dots, 2n$ each have row sum precisely equal to one. In other words, P is row-substochastic.

Notice that the graph $\mathcal{G}[P] = (\mathcal{V}, \mathcal{E}[P], P)$ has $2n$ nodes, with $\mathcal{V} = \{1, \dots, 2n\}$. The node subset $\mathcal{V}_1 = \{v_1, \dots, v_n\}$ contains node v_i which is associated with individual i 's private opinion y_i , $i \in \mathcal{I}$. The node subset $\mathcal{V}_2 = \{v_{n+1}, \dots, v_{2n}\}$ contains node v_{n+i} which is associated with individual i 's expressed opinion \hat{y}_i , $i \in \mathcal{I}$. Define the following two subgraphs; $\mathcal{G}_1 = (\mathcal{V}_1, \mathcal{E}[P_{11}], P_{11})$ and $\mathcal{G}_2 = (\mathcal{V}_2, \mathcal{E}[P_{22}], P_{22})$. The edge set of $\mathcal{G}[P]$ can be divided as follows

$$\begin{aligned} \mathcal{E}_{11} &= \mathcal{E} \begin{bmatrix} P_{11} & \mathbf{0}_{n \times n} \\ \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} \end{bmatrix}, & \mathcal{E}_{12} &= \mathcal{E} \begin{bmatrix} \mathbf{0}_{n \times n} & P_{12} \\ \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} \end{bmatrix}, \\ \mathcal{E}_{21} &= \mathcal{E} \begin{bmatrix} \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} \\ P_{21} & \mathbf{0}_{n \times n} \end{bmatrix}, & \mathcal{E}_{22} &= \mathcal{E} \begin{bmatrix} \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} \\ \mathbf{0}_{n \times n} & P_{22} \end{bmatrix}, \end{aligned}$$

In other words, \mathcal{E}_{11} contains only edges between nodes in \mathcal{V}_1 and \mathcal{E}_{22} contains only

edges between nodes in \mathcal{V}_2 . The edge set \mathcal{E}_{12} contains only edges from nodes in \mathcal{V}_2 to nodes in \mathcal{V}_1 , while the edge set \mathcal{E}_{21} contains only edges from nodes in \mathcal{V}_1 to nodes in \mathcal{V}_2 . Clearly $\mathcal{E}[\mathbf{P}] = \mathcal{E}_{11} \cup \mathcal{E}_{12} \cup \mathcal{E}_{21} \cup \mathcal{E}_{22}$. It will now be shown that $\mathcal{G}[\mathbf{P}]$ is strongly connected and aperiodic, which in turn establishes that \mathbf{P} is primitive.

Since the diagonal entries of $\mathbf{\Lambda}, \mathbf{\Phi}$ are strictly positive, it is obvious that $\mathbf{P}_{11} = \mathbf{\Lambda}(\widehat{\mathbf{W}} + \widehat{\mathbf{W}}\mathbf{\Phi}) \sim \mathbf{W}$. Because \mathbf{W} is primitive, it follows that \mathbf{P}_{11} is primitive, which in turn implies that \mathcal{G}_1 is strongly connected and aperiodic. Similarly, the edges of \mathcal{G}_2 are $\mathcal{E}[\mathbf{P}_{22}]$. Because $\mathbf{I}_n - \mathbf{\Phi}$ has strictly positive diagonal entries, one concludes that $\mathbf{P}_{22} = (\mathbf{I}_n - \mathbf{\Phi}) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \sim \mathcal{G}[\mathbf{1}_n \mathbf{1}_n^\top]$. From the fact that $\mathcal{G}[\mathbf{1}_n \mathbf{1}_n^\top]$ is a complete graph, it follows that \mathcal{G}_2 is a complete graph. Since \mathcal{G}_1 and \mathcal{G}_2 are both, separately, strongly connected, then if there exists 1) an edge from any node in \mathcal{V}_1 to any node \mathcal{V}_2 , and 2) an edge from any node in \mathcal{V}_2 to any node in \mathcal{V}_1 , one can conclude that the graph $\mathcal{G}[\mathbf{P}]$ is strongly connected. It suffices to show that $\mathcal{E}_{12} \neq \emptyset$ and $\mathcal{E}_{21} \neq \emptyset$, i.e., $\mathbf{P}_{21} = \mathbf{\Phi} \neq \mathbf{0}_{n \times n}$ and $\mathbf{P}_{12} = \mathbf{\Lambda} \widehat{\mathbf{W}} (\mathbf{I}_n - \mathbf{\Phi}) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \neq \mathbf{0}_{n \times n}$. Since $\mathbf{\Phi}$ has strictly positive diagonal entries, this proves that $\mathcal{E}_{12} \neq \emptyset$. From the fact that $\mathbf{I}_n - \mathbf{\Phi}$ has strictly positive diagonal entries, and because $\widehat{\mathbf{W}}$ is irreducible, it follows that $\mathbf{P}_{12} \sim \mathbf{1}_n \mathbf{1}_n^\top$. This shows that $\mathcal{E}_{21} \neq \emptyset$. It has therefore been proved that $\mathcal{G}[\mathbf{P}]$ is strongly connected and aperiodic, which also proves that \mathbf{P} is primitive. The arguments below Eq. (3.30) led to the conclusion that at least one row of \mathbf{P} has row sum strictly less than one. Lemma A.1 is then used to conclude that $\rho(\mathbf{P}) < 1$, i.e. all eigenvalues of \mathbf{P} are inside the unit circle. This completes the proof. \square

3.5.2 Proof of Lemma 3.2

First, observe that $\mathbf{Q}^{-1} = (\mathbf{I}_{2n} - \mathbf{P})^{-1}$. Because $\rho(\mathbf{P}) < 1$, it follows that \mathbf{P} is a bounded operator. That is, there exists a norm $\|\cdot\|'$ such that $\|\mathbf{P}\|' < 1$ (see [Bernstein, 2009, Corollary 9.3.4]). This implies that $\lim_{k \rightarrow \infty} \mathbf{P}^k = \mathbf{0}_{2n \times 2n}$. The Neumann series yields $\mathbf{Q}^{-1} = (\mathbf{I}_{2n} - \mathbf{P})^{-1} = \sum_{k=0}^{\infty} \mathbf{P}^k$. It was concluded in Section 3.5.1 above that $\mathcal{G}[\mathbf{P}]$ is strongly connected and aperiodic. Lemma 2.1 indicates that \mathbf{P} is primitive. This implies, from the Neumann series evaluation, that $\mathbf{Q}^{-1} > 0$. Next, it will be shown \mathbf{Q}_{11} , \mathbf{Q}_{22} and $\mathbf{Q}_{22} - \mathbf{Q}_{21} \mathbf{Q}_{11}^{-1} \mathbf{Q}_{12}$ are all invertible, which will allow Lemma A.2 to be used to express \mathbf{Q}^{-1} in the form of Eq. (3.11).

Recall that $\mathbf{Q}_{11} = \mathbf{I}_n - \mathbf{P}_{11}$ and $\mathbf{Q}_{22} = \mathbf{I}_n - \mathbf{P}_{22}$. One can verify that under Assumption 3.1, \mathbf{P}_{11} and \mathbf{P}_{22} are both primitive because $\mathcal{G}_1[\mathbf{P}_{11}]$ and $\mathcal{G}_2[\mathbf{P}_{22}]$ are both strongly connected and aperiodic, and $\rho(\mathbf{P}_{11}), \rho(\mathbf{P}_{22}) < 1$. Using precisely the same method as above, but which is omitted here, one can show that \mathbf{Q}_{11} and \mathbf{Q}_{22} are invertible, and that their inverses satisfy $\mathbf{Q}_{11}^{-1} > 0$ and $\mathbf{Q}_{22}^{-1} > 0$.

In order to prove that $\mathbf{Q}_{22} - \mathbf{Q}_{21} \mathbf{Q}_{11}^{-1} \mathbf{Q}_{12}$ is invertible, the matrix $\mathbf{S} = -\mathbf{Q}_{22}^{-1} \mathbf{Q}_{21}$ is first shown to be a positive, row-stochastic matrix. Since it was just established that \mathbf{Q}_{22}^{-1} is positive, it follows from the fact that $\mathbf{\Phi} = \text{diag}(\phi_i)$ is a diagonal matrix with positive diagonal entries, that $\mathbf{S} = \mathbf{Q}_{22}^{-1} \mathbf{P}_{21} = \mathbf{Q}_{22}^{-1} \mathbf{\Phi} > 0$. To prove that \mathbf{S} is row-stochastic, first note that \mathbf{S} is invertible because $\det(\mathbf{S}) = \det(\mathbf{Q}_{22}^{-1}) \det(\mathbf{\Phi}) \neq 0$

(because $\phi_i \in (0, 1), \forall i \Rightarrow \det(\Phi) \neq 0$). Observe that

$$S = \left(I_n - (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right)^{-1} \Phi = \left(\Phi^{-1} - \Phi^{-1} (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right)^{-1}, \quad (3.31)$$

because $(AB)^{-1} = B^{-1}A^{-1}$ for any invertible A, B . Observe that

$$S^{-1} \mathbf{1}_n = \left(\Phi^{-1} - \Phi^{-1} (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right) \mathbf{1}_n = \mathbf{1}_n - \Phi^{-1} \mathbf{1}_n + \Phi^{-1} \mathbf{1}_n = \mathbf{1}_n.$$

It then follows that $SS^{-1} \mathbf{1}_n = S \mathbf{1}_n \Leftrightarrow S \mathbf{1}_n = \mathbf{1}_n$, i.e. S has row sums equal to one. The proof that S is positive and row-stochastic is thus complete.

One now turns to proving that $T = Q_{11} - Q_{12} Q_{22}^{-1} Q_{21}$ is invertible, and because Q_{11} is invertible, this is equivalent to proving that $Q_{22} - Q_{21} Q_{11}^{-1} Q_{12}$ is invertible, see [Bernstein, 2009, pg. 108–109]. First, note that $S = -Q_{22}^{-1} Q_{21}$, and $-Q_{12} = P_{12} = \Lambda \widehat{W} (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n}$, and $\Lambda(\widehat{W} + \widehat{W}\Phi)$ are all nonnegative matrices. One can thus write $T = I_n - U$ where $U = \Lambda(\widehat{W} + \widehat{W}\Phi) + P_{12}S$ is a nonnegative matrix. Observe that

$$\begin{aligned} U \mathbf{1}_n &= \left(\Lambda(\widehat{W} + \widehat{W}\Phi) + \Lambda \widehat{W} (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} S \right) \mathbf{1}_n \\ &= \Lambda(\widehat{W} + \widehat{W}\Phi) \mathbf{1}_n + \left(\Lambda \widehat{W} (I_n - \Phi) \right) \mathbf{1}_n = \Lambda \mathbf{1}_n. \end{aligned}$$

where the last equality was obtained by recalling that $\widehat{W} + \widetilde{W} = W$ is a row-stochastic matrix, i.e. $(\widehat{W} + \widetilde{W}) \mathbf{1}_n = \mathbf{1}_n$. The above computation proves that the i^{th} row of U sums to $\lambda_i \in (0, 1)$ (see Assumption 3.1). In other words, U is positive and strictly row-substochastic, which implies that $\|U\|_\infty < 1$, and this in turn implies that $\rho(U) < 1$. Again, utilisation of the Neumann series shows that $T^{-1} = (I - U)^{-1} = \sum_{k=0}^{\infty} U^k$ converges to a finite limit. That is, T is invertible. Moreover, because it was shown in Section 3.5.1 that $\mathcal{G}[\Lambda(\widehat{W} + \widehat{W}\Phi)]$ is strongly connected and aperiodic, one can readily verify that $\mathcal{G}[U]$ is strongly connected and aperiodic. It follows that U is primitive, i.e. $\exists k \in \mathbb{N} : U^k > 0$. This primitivity property implies that $T^{-1} = \sum_{k=0}^{\infty} U^k > 0$. It follows that $R = T^{-1}(I_n - \Lambda)$ is positive, because $I_n - \Lambda$ is a diagonal matrix with positive diagonal entries. Finally, one can verify that R is row-stochastic with the following computation: $T \mathbf{1}_n = (I_n - U) \mathbf{1}_n = (I_n - \Lambda) \mathbf{1}_n \Rightarrow R \mathbf{1}_n = T^{-1}(I_n - \Lambda) \mathbf{1}_n = T^{-1} T \mathbf{1}_n = \mathbf{1}_n$. This completes the proof of Lemma 3.2. \square

3.5.3 Proof of Lemma 3.3

Note that the right hand side equality of Eq. (3.13) and Eq. (3.14), respectively, hold because the initial conditions are $\hat{y}(0) = y(0)$, as stated below Eq. (3.5). The proof will use Eq. (3.7) for simplicity, which can be written as

$$x(k+1) = Px(k) + Bx(0), \quad (3.32)$$

where $\mathbf{x}(k) = [\mathbf{x}_1(k)^\top, \mathbf{x}_2(k)^\top]^\top$, $k = 1, 2, \dots, \infty$, and with

$$\mathbf{B} = \begin{bmatrix} \mathbf{I}_n - \mathbf{\Lambda} & \mathbf{0}_{n \times n} \\ \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} \end{bmatrix}.$$

For $k = 1$, and as stated below Eq. (3.7), one has $\mathbf{x}_2(1) = \mathbf{x}_1(0)$ and $\mathbf{x}_1(1) = \mathbf{\Lambda}\mathbf{W}\mathbf{x}_1(0) + (\mathbf{I}_n - \mathbf{\Lambda})\mathbf{x}_1(0)$. In other words, $\mathbf{x}(1) = \mathbf{M}_1\mathbf{x}(0)$ where

$$\mathbf{M}_1 = \begin{bmatrix} \mathbf{\Lambda}\mathbf{W} + (\mathbf{I}_n - \mathbf{\Lambda}) & \mathbf{0}_{n \times n} \\ \mathbf{I}_n & \mathbf{0}_{n \times n} \end{bmatrix} \geq 0,$$

Verify that $\mathbf{M}_1\mathbf{1}_{2n} = \mathbf{1}_{2n}$. This indicates that each entry of $\mathbf{x}(1)$ is a convex combination of the entries of $\mathbf{x}_1(0)$. It follows that Eq. (3.13) and Eq. (3.14) hold for $k = 1$. The solution of Eq. (3.32), for $k > 1$, can be obtained in the follow recursive form:

$$\begin{aligned} \mathbf{x}(2) &= \mathbf{P}\mathbf{x}(1) + \mathbf{B}\mathbf{x}(0) = (\mathbf{P}\mathbf{M}_1 + \mathbf{B})\mathbf{x}(0) := \mathbf{M}_2\mathbf{x}(0), \\ &\vdots \\ \mathbf{x}(k-1) &= \mathbf{P}\mathbf{x}(k-2) + \mathbf{B}\mathbf{x}(0) = (\mathbf{P}\mathbf{M}_{k-2} + \mathbf{B})\mathbf{x}(1) := \mathbf{M}_{k-1}\mathbf{x}(0), \\ \mathbf{x}(k) &= \mathbf{P}\mathbf{x}(k-1) + \mathbf{B}\mathbf{x}(0) = (\mathbf{P}\mathbf{M}_{k-1} + \mathbf{B})\mathbf{x}(0) := \mathbf{M}_k\mathbf{x}(0). \end{aligned}$$

The matrices $\mathbf{M}_2, \mathbf{M}_2, \dots, \mathbf{M}_k$ are now shown to be row-stochastic. First, observe that

$$(\mathbf{P} + \mathbf{B})\mathbf{1}_{2n} = \mathbf{P}\mathbf{1}_{2n} + \mathbf{B}\mathbf{1}_{2n} = \begin{bmatrix} \mathbf{\Lambda}\mathbf{1}_n \\ \mathbf{1}_n \end{bmatrix} + \begin{bmatrix} (\mathbf{I} - \mathbf{\Lambda})\mathbf{1}_n \\ \mathbf{0}_n \end{bmatrix} = \begin{bmatrix} \mathbf{1}_n \\ \mathbf{1}_n \end{bmatrix} = \mathbf{1}_{2n},$$

where Eq. (3.30) has been used. In other words, $\mathbf{P} + \mathbf{B}$ is row-stochastic. Observe then that, for any row-stochastic matrix \mathbf{A} , $\mathbf{P}\mathbf{A} + \mathbf{B}$ is row-stochastic as well, because $(\mathbf{P}\mathbf{A} + \mathbf{B})\mathbf{1}_{2n} = (\mathbf{P} + \mathbf{B})\mathbf{1}_{2n}$. It was already shown that \mathbf{M}_1 was row-stochastic. It follows, by induction, that \mathbf{M}_k is row-stochastic for all $k \geq 1$, which, after recalling that $\mathbf{y}(t) = \mathbf{x}_1(k)$ and $\hat{\mathbf{y}}(t-1) = \mathbf{x}_2(k)$, proves the assertion in Lemma 3.3. \square

3.5.4 Simulation Counter-Example

The two inequalities Eq. (3.16) and Eq. (3.15) constitute a stronger form of nonexpansiveness which exists in other weight averaging algorithms such as the DeGroot or Hegselmann–Krause model (see Corollary 2.1). For the counter-example, consider a network of 4 individuals, with $\mathbf{y}(0) = [1, 0, 0, 0]^\top$, $\phi_1 = 0.8, \phi_2 = 0.03, \phi_3 = 0.02, \phi_4 = 0.05$, and $\lambda_1 = 0.1, \lambda_2 = 0.2, \lambda_3 = 0.12, \lambda_4 = 0.15$. The influence matrix is

$$\mathbf{W} = \begin{bmatrix} 0.2 & 0.2 & 0.3 & 0.3 \\ 0.2 & 0.3 & 0.2 & 0.3 \\ 0.1 & 0.3 & 0.4 & 0.2 \\ 0.2 & 0.2 & 0.1 & 0.5 \end{bmatrix}. \quad (3.33)$$

The trajectory of the system is shown in Fig. 3.6. Clearly $\max_{i \in \mathcal{I}} y_i(t)$ dips at $t = 1$ and then increases again for $t > 1$, which contradicts Eq. (3.15). A counter-example

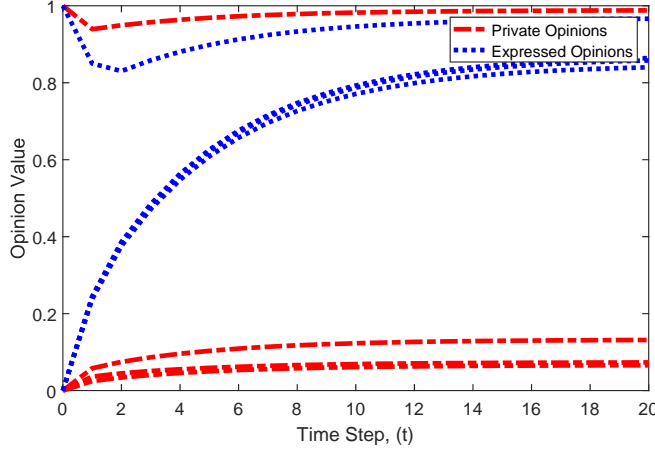


Figure 3.6: Evolution of opinions for counter-example to nonexpansive dynamics.

to Eq. (3.16) is obtained simply by setting $\mathbf{y}(0) = [-1, 0, 0, 0]^\top$.

3.5.5 Proof of Corollary 3.2

Recall the definition of the performance function V in Appendix A.1.1. From Theorem 3.1, one has that $V(\hat{\mathbf{y}}^*) = V(\mathbf{S}\mathbf{y}^*) \leq \tau(\mathbf{S})V(\mathbf{y}^*)$, which implies that there holds $V(\hat{\mathbf{y}}^*)/\tau(\mathbf{S}) \leq V(\mathbf{y}^*)$. Thus, Eq. (3.21) can be proved by showing that $\tau(\mathbf{S}) \leq \kappa(\boldsymbol{\phi})$. Recall that \mathbf{Q}_{22}^{-1} can be expressed using the Neumann series as $\mathbf{Q}_{22}^{-1} = (\mathbf{I} - \mathbf{P}_{22})^{-1} = \sum_{k=0}^{\infty} \mathbf{P}_{22}^k$. Since $\mathbf{P}_{22} = (\mathbf{I}_n - \boldsymbol{\Phi}) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n}$ and $\mathbf{Q}_{21} = -\boldsymbol{\Phi}$, it follows that

$$\mathbf{S} = -\mathbf{Q}_{22}^{-1}\mathbf{Q}_{21} = \sum_{k=0}^{\infty} \left[(\mathbf{I}_n - \boldsymbol{\Phi}) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right]^k \boldsymbol{\Phi} = \boldsymbol{\Phi} + \sum_{k=1}^{\infty} \left[(\mathbf{I}_n - \boldsymbol{\Phi}) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right]^k \boldsymbol{\Phi}.$$

For convenience, let $\underline{a} = \min_{i,j} a_{ij}$ denote the smallest element of a matrix \mathbf{A} . From the definition of τ (see Eq. (A.9)), it follows that, for any positive $\mathbf{A} \in \mathbb{R}^n$, there holds

$$\tau(\mathbf{A}) \leq 1 - n\underline{a} \quad (3.34)$$

Define $\mathbf{H} = \sum_{k=1}^{\infty} \left[(\mathbf{I}_n - \boldsymbol{\Phi}) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right]^k \boldsymbol{\Phi}$ and verify easily that $\mathbf{H} > 0$. Observe that there holds $\underline{s} = \underline{h}$ because $\mathbf{S} = \boldsymbol{\Phi} + \mathbf{H}$ has the same offdiagonal entries as \mathbf{H} , and the i^{th} diagonal entry of \mathbf{S} is greater than that of \mathbf{H} by ϕ_i . Since $\mathbf{S}, \mathbf{H} > 0$, it follows from Eq. (3.34) that $\tau(\mathbf{S}), \tau(\mathbf{H}) \leq 1 - n\underline{h}$, and thus the matrix \mathbf{H} will now be analysed.

Consider a matrix $A \in \mathbb{R}^{n \times n}$, with a_i denoting the i^{th} column. Observe that

$$\begin{aligned} (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} A &= \frac{1}{n} \begin{bmatrix} (1 - \phi_1) \mathbf{1}^\top \\ \vdots \\ (1 - \phi_n) \mathbf{1}^\top \end{bmatrix} [a_1 \ \cdots \ a_n] \\ &= \frac{1}{n} \begin{bmatrix} (1 - \phi_1) \sum_{j=1}^n a_{1j} & \cdots & (1 - \phi_1) \sum_{j=1}^n a_{nj} \\ \vdots & \ddots & \vdots \\ (1 - \phi_n) \sum_{j=1}^n a_{1j} & \cdots & (1 - \phi_n) \sum_{j=1}^n a_{nj} \end{bmatrix}. \end{aligned}$$

From this, it is straightforward to compute that

$$\left[(I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right]^2 = \frac{1}{n^2} \begin{bmatrix} (1 - \phi_1) \sum_{j=1}^n (1 - \phi_j) & \cdots & (1 - \phi_1) \sum_{j=1}^n (1 - \phi_j) \\ \vdots & \ddots & \vdots \\ (1 - \phi_n) \sum_{j=1}^n (1 - \phi_j) & \cdots & (1 - \phi_n) \sum_{j=1}^n (1 - \phi_j) \end{bmatrix}.$$

Similarly, by noting that $\sum_{i=1}^n \sum_{j=1}^n a_i b_j = \left(\sum_{i=1}^n a_i \right) \sum_{j=1}^n b_j = \sum_{i=1}^n a_i \left(\sum_{j=1}^n b_j \right)$, it follows that

$$\begin{aligned} \left[(I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right]^3 &= (I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \left[(I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right]^2 \\ &= \frac{1}{n^3} \begin{bmatrix} (1 - \phi_1) \sum_{j=1}^n \sum_{k=1}^n (1 - \phi_j)(1 - \phi_k) & \cdots & (1 - \phi_1) \sum_{j=1}^n \sum_{k=1}^n (1 - \phi_j)(1 - \phi_k) \\ \vdots & \ddots & \vdots \\ (1 - \phi_n) \sum_{j=1}^n \sum_{k=1}^n (1 - \phi_j)(1 - \phi_k) & \cdots & (1 - \phi_n) \sum_{j=1}^n \sum_{k=1}^n (1 - \phi_j)(1 - \phi_k) \end{bmatrix} \end{aligned}$$

By recursion, one obtains that the $(i, j)^{\text{th}}$ entry of $\left[(I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right]^k$ is given by

$$\frac{1}{n^k} (1 - \phi_i) \underbrace{\left[\sum_{p_1=1}^n \sum_{p_2=1}^n \cdots \sum_{p_{k-1}=1}^n (1 - \phi_{p_1})(1 - \phi_{p_2}) \cdots (1 - \phi_{p_{k-1}}) \right]}_{k-1 \text{ summation terms}}. \quad (3.35)$$

This is obtained using $\sum_{i=1}^n \sum_{j=1}^n a_i b_j = \left(\sum_{i=1}^n a_i \right) \sum_{j=1}^n b_j = \sum_{i=1}^n a_i \left(\sum_{j=1}^n b_j \right)$, except now operated on $k - 1$ summation terms. Next, define $Z^k = \left[(I_n - \Phi) \frac{\mathbf{1}_n \mathbf{1}_n^\top}{n} \right]^k \Phi$, and from Eq. (3.35), computations show that the $(i, j)^{\text{th}}$ element of Z^k is given by

$$z_{ij}(k) = \frac{1}{n^k} (1 - \phi_i) \phi_j \underbrace{\left[\sum_{p_1=1}^n \sum_{p_2=1}^n \cdots \sum_{p_{k-1}=1}^n (1 - \phi_{p_1})(1 - \phi_{p_2}) \cdots (1 - \phi_{p_{k-1}}) \right]}_{k-1 \text{ summation terms}}. \quad (3.36)$$

It follows that the smallest element of \mathbf{Z}^k , denoted by $\underline{z}(k)$, is bounded as follows

$$\underline{z}(k) \geq \frac{1}{n^k} (1 - \phi_{\max}) \phi_{\min} \underbrace{\left[\sum_{p_1=1}^n \sum_{p_2=1}^n \cdots \sum_{p_{k-1}=1}^n (1 - \phi_{p_1})(1 - \phi_{p_2}) \cdots (1 - \phi_{p_{k-1}}) \right]}_{k-1 \text{ summation terms}}. \quad (3.37)$$

Observe that $1 - \phi_i \geq 1 - \phi_{\max}, \forall i \Rightarrow \sum_{a=1}^n 1 - \phi_a \geq n(1 - \phi_{\max})$. It follows that

$$\underline{z}(k) \geq \frac{1}{n^k} (1 - \phi_{\max}) \phi_{\min} (1 - \phi_{\max})^{k-1} n^{k-1} \geq \frac{1}{n} \phi_{\min} (1 - \phi_{\max})^k. \quad (3.38)$$

Since $\mathbf{H} = \sum_{k=1}^{\infty} \mathbf{Z}^k$, it follows that

$$\begin{aligned} \underline{h} &\geq \sum_{k=1}^{\infty} \underline{z}(k) \\ &\geq \frac{1}{n} \phi_{\min} \left[\sum_{k=1}^{\infty} (1 - \phi_{\max})^k \right] \\ &\geq \frac{1}{n} \phi_{\min} \left[\frac{1}{1 - (1 - \phi_{\max})} - 1 \right] \\ &\geq \frac{1}{n} \left[\frac{\phi_{\min}}{\phi_{\max}} (1 - \phi_{\max}) \right], \end{aligned}$$

where the third inequality is obtained by recalling that for any $r \in \mathbb{R}$ such that $|r| < 1$, the geometric series is $\sum_{k=0}^{\infty} r^k = \frac{1}{1-r} \Leftrightarrow \sum_{k=1}^{\infty} r^k = \frac{1}{1-r} - 1$, and $0 < 1 - \phi_{\max} < 1$. It follows from Eq. (3.34), and the arguments immediately below it, that

$$\tau(\mathbf{S}) \leq 1 - n\underline{h} = 1 - \frac{\phi_{\min}}{\phi_{\max}} (1 - \phi_{\max}) = \kappa(\boldsymbol{\phi}) \quad (3.39)$$

as in the corollary statement. Since $0 < \phi_{\min}/\phi_{\max} < 1$ and $0 < 1 - \phi_{\max} < 1$, one has $0 < \frac{\phi_{\min}}{\phi_{\max}} (1 - \phi_{\max}) = \kappa(\boldsymbol{\phi}) < 1$ and thus $\tau(\mathbf{S}) \leq \kappa(\boldsymbol{\phi})$ holds $\forall \phi_i \in (0, 1)$. \square

3.5.6 Simulation Influence Matrix

$$W = \begin{bmatrix} 0.538 & 0.035 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.005 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.422 \\ 0.351 & 0.168 & 0.042 & 0 & 0 & 0 & 0 & 0 & 0 & 0.439 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.224 & 0.154 & 0.173 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.449 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.067 & 0.414 & 0.073 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.447 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.311 & 0.310 & 0.330 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.049 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.496 & 0.047 & 0.305 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.152 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.231 & 0.328 & 0.220 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.221 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.602 & 0.001 & 0.193 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.203 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.191 & 0.332 & 0.256 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.221 \\ 0.486 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.044 & 0.362 & 0.108 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.280 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.362 & 0.126 & 0.232 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.074 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.424 & 0.343 & 0.159 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.336 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.317 & 0.029 & 0.318 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.190 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.199 & 0.315 & 0.297 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.300 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.542 & 0.082 & 0.077 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.402 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.034 & 0.172 & 0.393 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.252 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.312 & 0.253 & 0.182 & 0 \\ 0.267 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.212 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.182 & 0.339 & 0 \end{bmatrix}$$

The EPO Model's Connections with Social Psychology Concepts

Chapter Summary

This chapter will study the Expressed–Private Opinion (EPO) model first introduced in the previous Chapter 3. First, the model is used to revisit and re-examine Asch's seminal conformity experiments. Analytical calculations are used to establish the opinion evolution of the test individuals in the experiments. With the aid of simulations, it is shown that all of the different ways that individuals were recorded to have reacted in the experiments are captured and predicted by the proposed model. Then, *pluralistic ignorance* is investigated. This phenomenon considers how the perceived and true positions of the general population on a given topic can be completely different. Using extensive simulations, the role of stubborn extremists (termed zealots) in creating pluralistic ignorance is revealed.

4.1 Introduction

A number of works from the sociology and social psychology communities were presented in Section 3.1, introducing concepts such as how an individual may have different expressed and private opinions, and how individuals often face a pressure to conform to a group norm or standard. Continuing from that, two established results are further investigated.

4.1.1 Introduction to Asch's Experiments

The context in which Asch's seminal experiments on conformity to pressure is relevant to the proposed model was outlined in the introduction of Chapter 3. In addition to providing much of the motivation for the EPO model, it is also of great interest to revisit Asch's results using the model. This can be used to provide some validation for the model, while simultaneously gaining access to new insight via analysis from a networked, agent-based mathematical framework.

In order for the reader to fully appreciate and understand the results to be covered in this chapter, a brief overview of Asch's conformity experiments and its results are now given, and the reader is referred to [Asch, 1951] for full details on the results. In summary, the Asch experiments studied an individual's response to "two contradictory and irreconcilable forces" of (i) a clear and indisputable fact, and (ii) a unanimous majority of the others in the group who take positions opposing this fact.

In the first set of experiments (which is what this chapter will focus on), eight individuals (all male college students) are instructed to judge a series of line lengths. Of the eight individuals, one is in fact the test subject, and the other seven have been told a priori about what they should do. These other individuals have become referred to as "confederates" in later literature for convenience, and this thesis adopts this terminology. An example of the line length judging experiment, and the influence network, is shown in Figure 4.1. There are three lines of unequal length, and the group has open discussions concerning which one of the lines *A, B, C* is equal in length to the green line. Each individual is required to independently declare his choice. The group is required to conduct this experiment for twelve iterations in sequence (using different sets of lines for each experiment). For each iteration, there is an indisputable and obvious answer. Control groups without confederates were extremely successful (close to 100% of groups) in determining the correct answer, and so there is no issue of ambiguity in the correct answer.

At predetermined iterations, the seven confederates unanimously and confidently choose the *same wrong answer*. They stubbornly persist in this answer and even openly question the choice made by the eighth individual. The actions of the eighth, the test individual, are recorded to determine whether he also selects the wrong answer, or maintains his expressed belief for the correct answer in the face of pressure to conform to a unanimous majority. After the experiment, the test individual is interviewed to further understand his decisions and thoughts. These interviews were critical in Asch's classification of the individuals. In particular, Asch identified three broad types of test individuals, these being (i) *independent* individuals, (ii) *yielding* individuals who showed *distortion of judgment/perception*, and (iii) *yielding* individuals who showed *distortion of action*. The nature of each type of individual will be clarified in the sequel, when the Asch experiments are framed within the EPO model.

4.1.2 Pluralistic Ignorance in Social Networks

Although racial segregation still existed in United States in the 1960s, it turns out that there was a misconception among the population at large about the popularity of segregation. Specifically, [O'Gorman, 1975] found that at the time, segregation was not supported by the majority of white Americans, but *the majority of white Americans believed that segregation enjoyed majority support among white Americans*. This is a phenomenon known as *pluralistic ignorance*, in which there is a misconception about the true views of the population. Closely related to pluralistic ignorance is the social psychological phenomenon called "the spiral of silence" [Noelle-Neumann, 1993], in which an individual may cease to express his/her opinion if that individual percei-

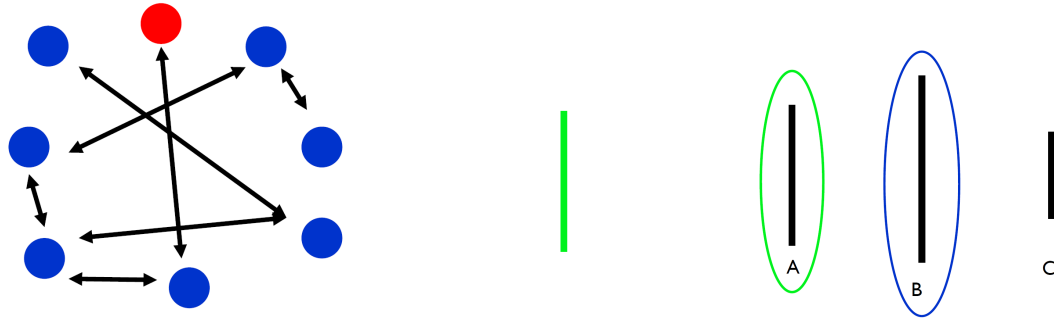


Figure 4.1: Example of the Asch experiment, with an illustrative influence network of eight individuals in the Asch. Here, the edges are bidirectional (for convenience only), indicating both individuals influence one another. The individuals are required to openly discuss the length of lines, and state their individual beliefs as to which one of A, B, C has the same length as the green line. Clearly A is equal in length to the green line. The confederates (seven blue individuals) unanimously express belief in the same wrong answer, e.g. B .

ves the general public to be shifting away from his/her current opinion. Conversely, individuals may become extremely vocal even when they hold an opinion in the minority, because they perceive the general public's opinions to be shifting toward their current position. The two phenomena have been studied together [Taylor, 1982]; pluralistic ignorance is typically seen as a reflection of the individual's *knowledge* of the current state of the social network, whereas the spiral of silence is an *act* from an individual based on predictions of the *future state* of the network (these predictions may be based on inaccurate information arising from pluralistic ignorance).

The psychological process of *preference falsification* was introduced in Section 1.2.3, and is now explained in further detail to aid in the discussion of the results to be presented in Section 4.3. The term is widely attributed to political and economic scientist Timur Kuran, as discussed in [Kuran, 1997, Chapter 1]. Preference falsification occurs when an individual misrepresents his/her genuine (i.e. private) wants and/or opinions under social pressure that is perceived by the individual. A politician may deliberately falsify his/her expressed preferences to garner the votes of the people. An individual may unconsciously falsify his/her preference under real or imaginary social pressure arising from a group situation. Often times, an individual will falsify his/her preference in order to please, or avoid conflict and confrontation with an individual(s) he/she likes or wishes to impress. From this point of view, Eq. (3.2) represents individual i falsifying i 's private (true) opinion, and specifically in the direction of the public opinion \hat{y}_{avg} . For the purposes of this thesis, no attempt will be made to differentiate whether there is deliberate or unconscious falsification; focus is instead placed on the fact that an individual *has acted to falsify his/her preference*. Moreover, and in the context of this thesis, pluralistic ignorance is a *measure of a group or individual's incorrect knowledge* of the opinions of the wider social network, whereas preference falsification is an *action*.

It is obvious why misconceptions about the views of the general population are,

in most circumstances, undesirable. It was already mentioned in Chapter 3 that large differences in the private and expressed opinions of a population can lead to violent and unexpected actions, such as the Arab Spring movement and the fall of the Soviet Union. Without a clear knowledge of what the population's positions are on a given topic, there is little hope of genuine discourse for divisive topics. Recent events have highlighted the dangers of large scale social media in the spreading of conspiracies and falsehoods [Del Vicario et al., 2016], and the use of hostile bot accounts on Twitter, Facebook, and other social media to influence and manipulate voters in recent elections [McCright and Dunlap, 2017; Calabresi, 2017; Rizoio et al., 2018]. Using simulations, Section 4.3 will show that just a few stubborn and resilient extremists (termed zealots) are capable of creating massive pluralistic ignorance in a large population. It turns out that some network types are more susceptible to disruption by zealots than others, and further analysis provides several methods for reducing the impact of the zealots.

4.1.3 Chapter Organization

The remainder of the chapter has the following structure. Section 4.2 studies the Asch experiments using the EPO model via analytical calculations and simulations. The phenomenon of pluralistic ignorance is investigated in Section 4.3, and concluding remarks are given in Section 4.4.

4.2 Investigation of Asch's Conformity Experiments

Since Asch's experiments deal with an individual's position on a statement which is provable to be true or false¹, the variables y_i, \hat{y}_i are referred to as "beliefs" *only in this section*, i.e. Section 4.2. For all other parts of the thesis, the use of "opinion" is maintained when referring to y_i, \hat{y}_i , for the sake of consistency.

On first observation, Asch's experiments presented individuals with a set of discrete choices, and thus the proposed model, with $y_i, \hat{y}_i \in \mathbb{R}$, is not applicable. However, the EPO model is in fact applicable for analysis purposes if the beliefs y_i, \hat{y}_i are framed correctly. Consider Figure 4.1 as the example. The topic of discussion is on the truth of the statement: "The green line is of the same length as line A." In this section, the beliefs y_i, \hat{y}_i are scaled to be in $[0, 1]$, in which $y_i = 1$ (respectively $y_i = 0$) implies individual i is maximally certain that A is of the same length as the green line (respectively maximally certain that A is not of the same length as the green line). Because line A being equal in length to the green line is an indisputable fact, it is expected for non-confederates to have $y_i(0)$ close to, if not equal to 1. This is supported in [Asch, 1951], where the control groups had no difficulty in obtaining the correct line length answer.

Without loss of generality, it is assumed that individual 1 is the test subject, with individuals $2, \dots, n$ being the confederates. The initial conditions are $\mathbf{y}(0) = \hat{\mathbf{y}}(0) =$

¹See Section 2.3.3 for the differences between opinion and belief as defined in this thesis.

$[1, 0, \dots, 0]^\top$, i.e. test individual 1 has selected the correct answer A , while the group (v_2, \dots, v_n) is unanimous in claiming that A is not correct. For $i = 2, \dots, n$ set $\lambda_i = 0$ and $\phi_i = 1$. That is, the confederates are maximally stubborn in their initial belief (since they were informed a priori by the experimenter to unanimously choose the wrong answer) and their private and expressed beliefs are equal, $y_i = \hat{y}_i$, because they have been told to be maximally resilient. The global public opinion updating rule Eq. (3.2) is used, since the network is small and thus $y_{\text{avg}}(t)$ is easily known for the test individual.

It should be noted here that in the experiment, Asch never assigned values of susceptibility λ_i , and resilience ϕ_i to the test individuals because the quantitatively measured data by Asch was the number of incorrect answers over 12 iterations and the behaviour of the individual being tested. However, based on his written description of individuals (including excerpts of the interviews), it was clear to the author of this thesis what the approximate range of values of the parameters λ_i, ϕ_i should be for each type of individual. (Some of these descriptions and excerpts will be provided immediately below). Also, the experiments did not attempt to determine the influence matrix W (at the time, influence network theory in the sense of DeGroot etc. had not yet been developed). The qualitative observations made in this section are invariant to the weights w_{ij} , and focus is instead placed on examining Asch's experimental results from the perspective of the EPO model. In the following Section 4.2.1, the impact of W (and in particular the weights w_{11}), and parameters ϕ_1 and λ_1 , are shown using analytic calculations.

Types of Individuals

Asch observed three broad types of individuals. In particular, he divided the test individuals as: (i) independent individuals, (ii) yielding individuals with distortion of judgment, and (iii) yielding individuals with distortion of action.

Independent individuals in [Asch, 1951] can be divided further into different subgroups depending on the reasoning behind their independence, but this will not be considered because this chapter focuses on only on the final outcome or observed result and not the reasons. Asch identified an independent individual as someone who was strongly confident that A was correct. This individual did not change his expressed belief, i.e. did not yield to the confederates' unanimous declaration that A was incorrect, despite the confederates insistently questioning the individual. Asch's descriptions indicate that the test individual is extremely stubborn (i.e. closed to influence) and confident his belief is correct, and is resilient to the group pressure. It is then obvious that one expects λ_1 to be close to zero and ϕ_1 to be close to one. Values of ϕ_1, λ_1 in this neighbourhood generate responses that are qualitatively the same at a high level; the differences lie in the exact values of the final opinions. This will become apparent in the following two sections.

Asch also identified *yielding* individuals, who could be divided into two groups. Those who experienced a distortion of *judgment/perception* either (i) lacked confidence, assumed the group was correct and thus concluded A was incorrect, or (ii)

did not realise he had been influenced by the group at all and changed his private belief to be certain that A was incorrect. This indicates that the individual is open to influence (i.e. not stubborn in $y_1(0) = 1$) and is highly affected by the group pressure (i.e. not resilient). One concludes that λ_1 is likely to be close to one, and ϕ_1 to be close to zero. As shown in the sequel, it turns out that the value of ϕ_1 plays only a minor role for such an individual because he is already extremely open to influence.

Other yielding individuals experienced a *distortion of action*. This type of individual, on being interviewed (and before being informed of the true nature of the experiment) stated that he remained privately certain that A was the correct answer, but suppressed his observations as to not publicly generate friction with the group. Such an individual has full *awareness of the difference* between the truth and the majority's position. This individual is closed to influence (i.e. stubborn) but not resilient, and it is predicted that such individuals will have λ_1 and ϕ_1 both close to zero.

4.2.1 Theoretical Analysis

This section will present theoretical calculations of Asch's experiments in the framework of the EPO model, and provide analytic expressions to describe how the test subject's beliefs y_1, \hat{y}_1 vary with W , $\lambda_1 \in [0, 1]$ and $\phi_1 \in [0, 1]$. Analysis will be conducted for arbitrary $n \geq 2$, to also investigate if the increasing of the majority size affects the belief evolution. The details of the confederates' parameters are given at the start of Section 4.2. It is assumed that $w_{11} \in (0, 1)$, which is mild, and simply implies that individual 1 considers his/her own belief during the discussions.

Because $\lambda_i = 0$ and $\phi_i = 1$ for all $i = 2, \dots, n$, one concludes from Eq. (3.1) and Eq. (3.2) that $y_i(t) = \hat{y}_i(t) = 0$ for all t . Recalling the initial conditions $\mathbf{y}(0) = [1, 0, \dots, 0]^\top$, it follows that individual 1's belief evolves as

$$\begin{bmatrix} y_1(t+1) \\ \hat{y}_1(t) \end{bmatrix} = \begin{bmatrix} \lambda_1 w_{11} & 0 \\ \phi_1 & \frac{1}{n}(1 - \phi_1) \end{bmatrix} \begin{bmatrix} y_1(t) \\ \hat{y}_1(t-1) \end{bmatrix} + \begin{bmatrix} 1 - \lambda_1 \\ 0 \end{bmatrix}. \quad (4.1)$$

From the fact that $n \geq 2$, $\lambda_1 \in [0, 1]$, $w_{11} \in (0, 1)$, and $\phi_1 \in [0, 1]$, it follows that

$$\mathbf{V} = \begin{bmatrix} \lambda_1 w_{11} & 0 \\ \phi_1 & \frac{1}{n}(1 - \phi_1) \end{bmatrix} \quad (4.2)$$

has eigenvalues inside the unit circle and thus the system in Eq. (4.1) converges to a unique equilibrium exponentially fast (this is consistent with the results in Section 3.3). This equilibrium is given by

$$\begin{bmatrix} \lim_{t \rightarrow \infty} y_1(t) \\ \lim_{t \rightarrow \infty} \hat{y}_1(t) \end{bmatrix} \triangleq \begin{bmatrix} y_1^* \\ \hat{y}_1^* \end{bmatrix} = (\mathbf{I}_2 - \mathbf{V})^{-1} \begin{bmatrix} 1 - \lambda_1 \\ 0 \end{bmatrix}, \quad (4.3)$$

where I_2 is the 2×2 identity matrix, and one obtains

$$(I_2 - V)^{-1} = \begin{bmatrix} (1 - \lambda_1 w_{11})^{-1} & 0 \\ \frac{n\phi_1}{(1 - \lambda_1 w_{11})(n - 1 + \phi_1)} & \frac{n}{1 - \phi_1} \end{bmatrix}.$$

It follows that the final beliefs are given as

$$y_1^* = \frac{1 - \lambda_1}{1 - \lambda_1 w_{11}} \quad (4.4)$$

$$\hat{y}_1^* = \frac{n\phi_1}{n - 1 + \phi_1} y_1^*. \quad (4.5)$$

From this, one concludes that the test subject's final private belief is dependent on his level of stubbornness in believing that A is the correct answer, i.e. λ_1 , and on his self-weight w_{11} , i.e. how much he trusts his own belief relative to the others in the group. Interestingly, y_1^* does not depend on individual 1's resilience ϕ_1 , though it must be noted that this is a special case when the other individuals are all confederates. In general, y_1^* will depend not only on ϕ_1 , but also the other ϕ_i (as discussed in Section 3.3). For simplicity, consider a natural selection of $w_{ii} = 1 - \lambda_i$ (see Section 2.3.2). As a result, one obtains that $y_1^* = (1 - \lambda_1)/(1 - \lambda_1(1 - \lambda_1))$. Examination of the function $f(\lambda_1) = (1 - \lambda_1)/(1 - \lambda_1(1 - \lambda_1))$, for $\lambda_1 \in [0, 1]$, reveals how the test subject's final private belief changes as a function of his openness to influence; the function $f(\lambda_1)$ is plotted in Fig. 4.2. Notice that $f(\lambda_1) = (1 - \lambda_1)/(1 - \lambda_1(1 - \lambda_1)) \geq 1 - \lambda_1$ for $\lambda_1 \in [0, 1]$ with equality if and only if $\lambda_1 = \{0, 1\}$. This implies that the test individual's final y_1^* will always be greater than his stubbornness $1 - \lambda_1$, except if he has $\lambda_1 = 0$ (maximally stubborn) or $\lambda_1 = 1$ (maximally open to influence).

Next, consider the final expressed belief, which is given as $\hat{y}_1^* = \frac{n\phi_1}{n - 1 + \phi_1} y_1^*$. The relative closeness of \hat{y}_1^* to y_1^* , as measured by \hat{y}_1^*/y_1^* , is determined by n and ϕ_1 . Define $g(\phi_1, n) = \frac{n\phi_1}{n - 1 + \phi_1}$. The function $g(\phi_1, n)$ is plotted in Fig. 4.3. Observe that $g(\phi_1, n) \geq \phi_1$ for any n , for all $\phi_1 \in [0, 1]$, and with equality if and only if $\phi_1 = \{0, 1\}$. This implies that the test individual's final expressed belief will always be closer to his final private belief than his resilience level. Most interestingly, observe that $g(\phi_1, n) \rightarrow \phi_1$ from above, as $n \rightarrow \infty$, but the difference between $g(\phi_1, n)$ and ϕ_1 when going from $n = 2$ to $n = 2 \times 2 = 4$ is much greater than the differences going from $n = 4$ to $n = 4 \times 2 = 8$. This may explain the observation in [Asch, 1951] that increasing the majority size did not produce a correspondingly larger distortion effect beyond majorities of three to four individuals, at least for test individuals with low λ_1 . That is, an increase in n does not produce a matching increase in distortion of the final expressed opinion from the final private opinion, represented as $\hat{y}_1^*/y_1^* = g(\phi_1, n) \rightarrow 1$ as $n \rightarrow \infty$.

Also of note is that for individuals with λ_1 close to one, y_1^* is already close to zero, and bounds \hat{y}_1^* from above. The *magnitude of the difference*, $|y_1^* - \hat{y}_1^*|$, only changes slightly as ϕ_1 is varied, which indicates that for individuals who yielded with distortion of judgment, the value of ϕ_1 plays only a minor role in the determining the absolute (as opposed to relative) difference between expressed and private beliefs. This is in

Table 4.1: The types of individuals and their associated parameters in simulation examples

	Susceptibility λ_1	Resilience ϕ_1
Independent	$\lambda_1 = 0.1$	$\phi_1 = 0.9$
Yielding, distortion of judgment	$\lambda_1 = 0.9$	$\phi_1 \in [0, 1]$
Yielding, distortion of action	$\lambda_1 = 0.1$	$\phi_1 = 0.1$

contrast to individuals with low susceptibility, where the behaviour of an individual can vary significantly by varying ϕ_1 from 1 to 0. The relation between ϕ_1, λ_1 and the type of test individual becomes clear in the following section, when simulations are used to illustrate each type of test individual.

4.2.2 Simulations

The Asch experiments are now simulated using the EPO model. It was mentioned in the last paragraph of Section 4.2 that \mathbf{W} was not identified by Asch during the experiments. The effect of w_{11} was studied theoretically in the previous Section 4.2.1. In these simulations an arbitrary \mathbf{W} is generated with weights w_{ij} sampled randomly from a uniform distribution and normalised to ensure $\sum_{j=1}^n w_{ij} = 1$. As above, individual 1 is set as the test subject, with individuals $2, \dots, 8$ being the 7 confederates in an $n = 8$ network. The initial conditions are $\mathbf{y}(0) = \hat{\mathbf{y}}(0) = [1, 0, \dots, 0]^\top$, while $\lambda_i = 0$ and $\phi_i = 1$ for $i = 2, \dots, 8$. Table 4.1 gives an overview of the three types of test individuals observed in the Asch experiments, and the values that the author assigned for the associated parameters of openness to influence λ_1 , and resilience to conformity ϕ_1 (see page 55 for a commentary on the types of individuals and the author assigned approximate values of λ_1, ϕ_1).

In the following plots of Fig. 4.5, 4.6 and 4.7, red lines correspond to test individual 1; the solid line shows the time evolution of his private belief $y_1(t)$ and the dotted line shows the time evolution of his expressed belief, $\hat{y}_1(t)$. The blue line represents the confederates v_2, v_3, \dots, v_8 , i.e. the seven other participants; since they are maximally resilient and closed to influence, $y_k(t) = \hat{y}_k(t) = 0$ for all t , and $k = 2, \dots, 8$.

Discussion of Simulation Results

Figure 4.5 shows the evolution of beliefs when the test individual is *independent*. It can be seen that both the private and expressed beliefs of v_1 are largely unaffected by the confederates' unanimous expressed belief and the pressure exerted by the group. Note that $\hat{y}_1^* < y_1^*$, which is also reported in [Asch, 1951]; despite expressing his belief that A is the correct answer, one independent test individual stated "You're *probably* right, but you may be wrong!", which might be seen as a concession towards the majority belief. There is also a small shift away from maximal certainty of $y_i = 1$, with $y_1^* \approx 0.93$; in [Asch, 1951], one independent test individual stated

I would follow my own view, though part of my reason would tell me that I might

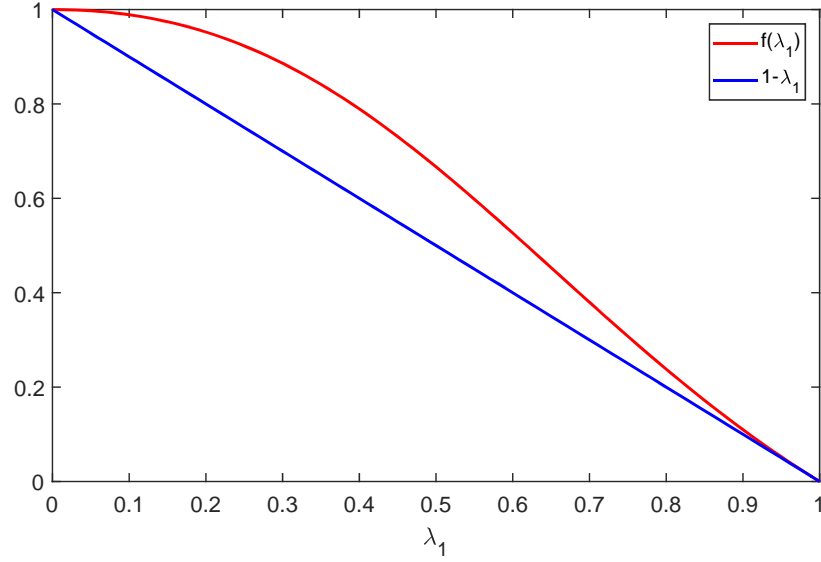


Figure 4.2: The function $f(\lambda_1)$ and $1 - \lambda_1$ plotted against λ_1 . The analytical calculations show that $y_1^* = f(\lambda_1)$, and thus the red line represents individual 1's final private belief as a function of his susceptibility to influence.

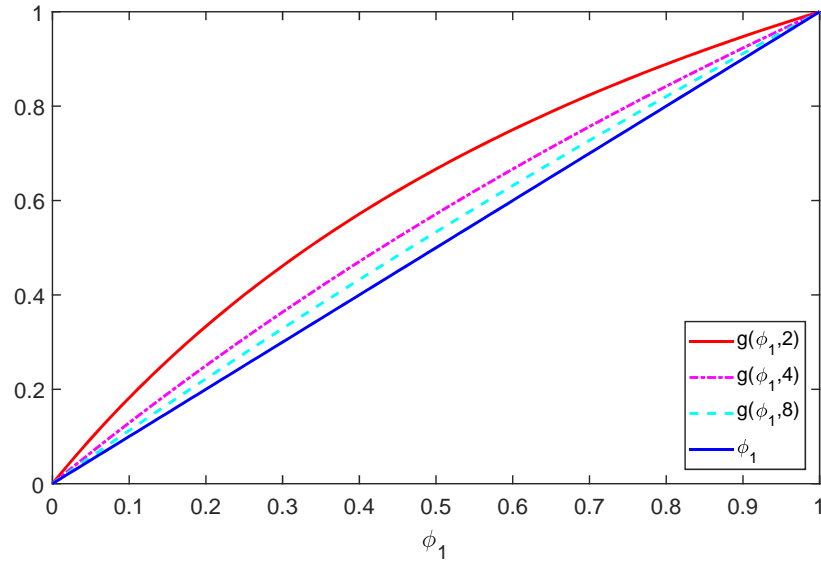


Figure 4.3: The function $g(\phi_1, n)$, with $n = 2, 4, 8$, plotted against ϕ_1 . The analytical calculations show that $\hat{y}_1^* = g(\phi_1, n)y_1^*$, and thus the plot shows how the test individual's final expressed opinion is changed from his final private opinion by his resilience ϕ_1 , and by n .

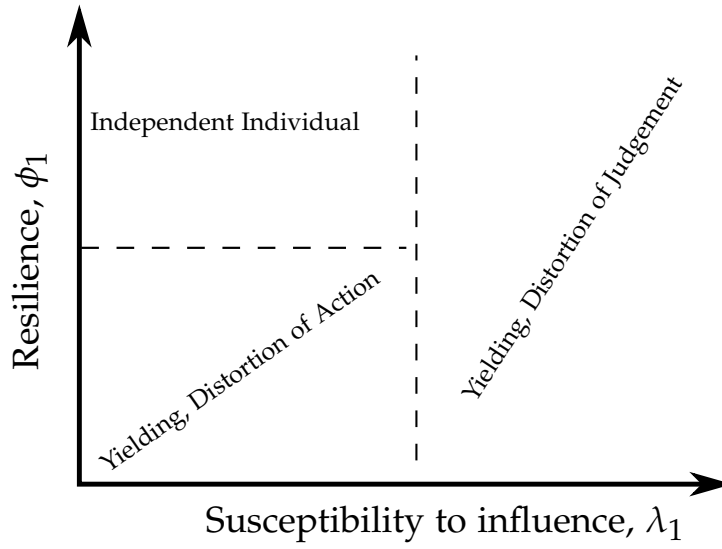


Figure 4.4: Behaviour of the test individual as a function of susceptibility to influence, λ_1 , and resilience to pressure, ϕ_1 .

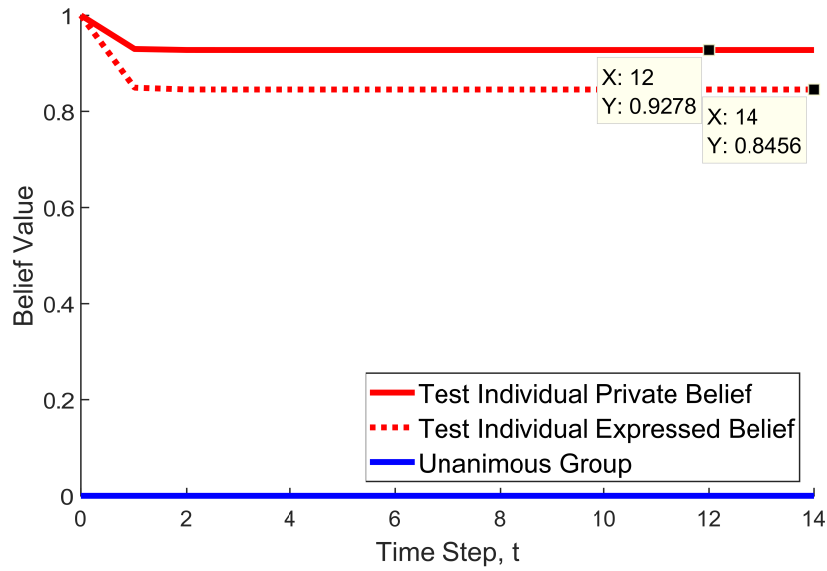


Figure 4.5: Evolution of beliefs for an *independent* individual, $\lambda_1 = 0.1, \phi_1 = 0.9$.

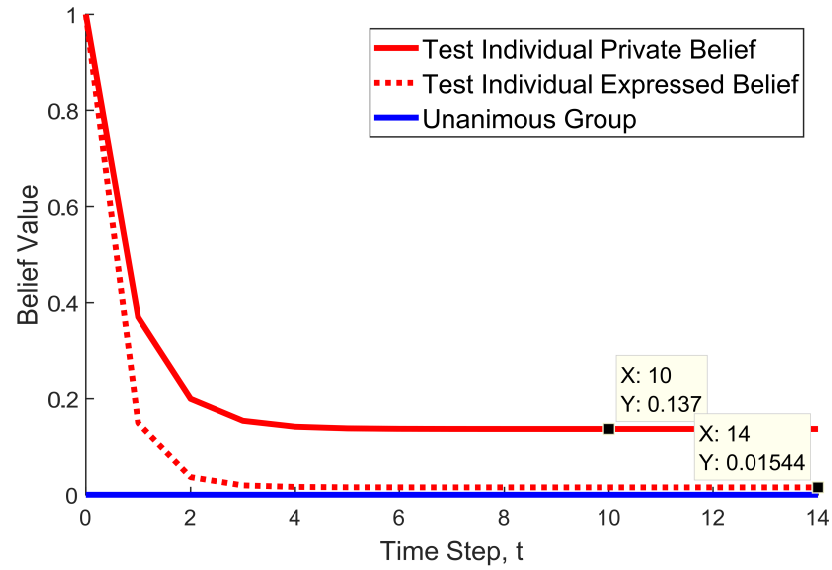


Figure 4.6: Evolution of beliefs for a *yielding* individual with *distortion of judgment/perception*, $\lambda_1 = 0.9, \phi_1 = 0.1$.

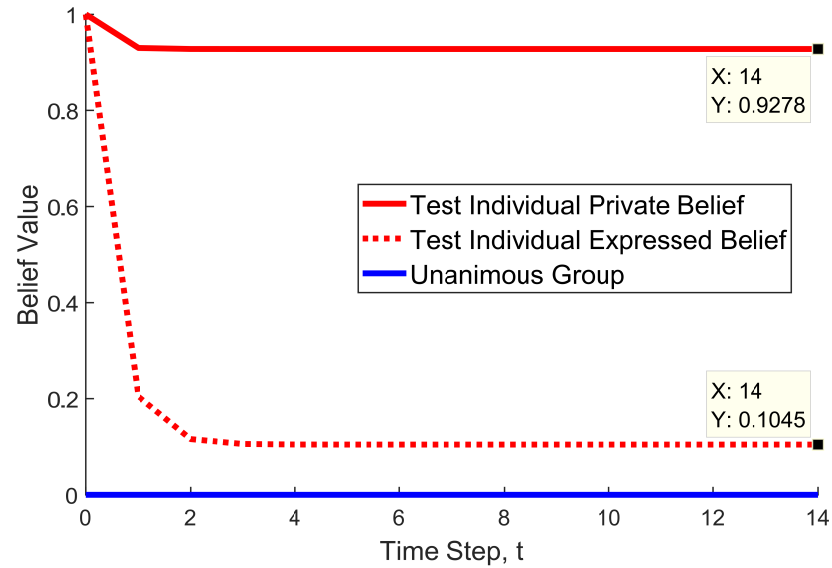


Figure 4.7: Evolution of beliefs for a *yielding* individual with *distortion of action*, $\lambda_1 = 0.1, \phi_1 = 0.1$.

be wrong.

Figure 4.6 shows the belief evolution of a yielding test individual who, under group pressure, exhibits *distortion of judgment/perception*. The figure shows that both y_1^* and \hat{y}_1^* are heavily influenced by the group pressure, and thus individual 1 is no longer privately certain that A is the correct answer. In other words, this individual is highly susceptible to interpersonal influence, and even his private view becomes affected by the majority. Of great interest is the evolution of beliefs observed in Fig. 4.7, which involves an experiment with a yielding test individual exhibiting *distortion of action*. According to Asch, Individual 1

yields because of an overmastering need to not appear different or inferior to others, because of an inability to tolerate the appearance of defectiveness in the eyes of the group ~[Asch, 1951].

In other words v_1 's expressed belief y_1^* is heavily distorted by the pressure to conform to the majority. However, this individual is still able to "conclude that they [themselves] are not wrong" [Asch, 1951], i.e. $y_i^* \approx 0.93$.

Other simulations with parameters of λ_1, ϕ_1 in the neighbourhood of those identified in Table 4.1 also display similar behaviour as shown in Fig. 4.5 to 4.7. This is captured by Fig. 4.4, which shows the type of individual as a function of λ_1 and ϕ_1 . From the theoretical and simulation results, it is clear that all three types of individuals can be predicted by the EPO model using pairs of parameters λ_i, ϕ_i , and that their behaviour in the EPO model accurately captures the empirical findings of Asch's experiments. This provides a measure of validation for the EPO model, while at the same time, provides a useful way for analysing the Asch experiments (and the many subsequent works derived from the paper) in a mathematical context. It is noted here that in [Friedkin and Johnsen, 2011], Noah E. Friedkin and Eugene Johnsen also applied the Friedkin–Johnsen model to the Asch experiments. However, their model was not able to capture all of the types of individuals reported, and specifically was not able to differentiate between the two different types of yielding individuals. This is because the Friedkin–Johnsen model does not assume that each individual has a separate private and expressed belief.

4.3 A Few Zealots Can Create Pluralistic Ignorance

In this section, a number of simulations of the EPO model are used to show how pluralistic ignorance and preference falsification can occur in large networks as a result of only a few zealots (the definition of which will be provided in the sequel). The purpose of the simulations is not to provide a complete characterisation of the pluralistic ignorance phenomenon and how the opinion trajectories change with network type, distributions of $\phi_i, \lambda_i, y_i(0)$, etc. This requires an exhaustive and comprehensive study, which is left as possible future work. Instead, these simulations are intended to demonstrate the use of the model in revealing and examining the causes of pluralistic ignorance. This section also explores the mechanics that drive this phenomenon

under the EPO model. This is done by studying the EPO model on both small-world and scale-free networks. In addition, both the local and global public opinion variants are investigated. It turns out that for scale-free networks, there is a major difference in the level of pluralistic ignorance depending on whether the local or global public opinion model is used, whereas there is less of a difference for scale-free networks. These differences are then examined to help explain how the impact of the zealots might be reduced

Several slightly different definitions of pluralistic ignorance exist. This thesis adopts the broader definition, as detailed by Robert K. Merton in [Merton, 1968]. In particular, pluralistic ignorance refers to the situation whereby individuals hold an incorrect assumption about the thoughts, feelings and/or behaviour of other people. The most common situation, and the situation dealt with in this thesis, is where a group of individuals² *underestimate* the number of others who hold similar opinions³. Specifically, the individuals hold (private) opinions similar to that of the majority but think that they are in the minority because of what is observed. The definition also includes the situation where individuals incorrectly think they hold (private) opinions similar to the majority, but are in fact holding opinions similar to the minority.

To begin, the base simulation set-up is explained. Then, the specifics of the small-world and scale-free network simulations are explained, and results presented. In a separate section, discussions are provided on the results.

4.3.1 Base Simulation Set-Up

The base set-up for the simulations is now described; specific changes to this set-up are detailed in the relevant latter sections. It is of interest to investigate the effects of increasing the number of zealots within a network of $n = 200$ individuals. The opinions are initialised in the interval $[0, 1]$. From Lemma 3.3, all opinions will stay in the interval $[0, 1]$ for all time. The definition of a zealot is now introduced.

Definition 4.1 (Zealot). *An individual i is said to be a zealot⁴ if $y_i(0) = 1$ and $\lambda_i = 0.001$ and $\phi_i = 0.999$.*

According to Eq. (3.1) and Eq. (3.2), a zealot holds (for practical purposes) an opinion at one extreme of the interval $[0, 1]$, and is (for practical purposes) maximally stubborn and maximally resilient to pressures to conform to the public opinion. Zealots may be partisan media sources, fake news websites, leader figures, or individual fanatics [McCright and Dunlap, 2017]. The fraction of zealots in a network of n individuals is defined as f_z so that $f_z n$ is the total number of zealots, and $f_z(1 - n)$

²The definition in [O’Gorman, 1975] is given for two or more individuals.

³Note that the literature also refers to “the same opinion”, because the opinions considered in the social science literature are often discrete variables. Since this thesis deals with opinions as real values on a continuous interval, the word “similar” is used to mean opinions which are approximately of the same value.

⁴The choice of values for λ_i and ϕ_i are arbitrary, and selected only for convenience in running the simulations while ensuring Assumption 3.1 holds. The results and conclusions drawn in this section will hold if zealots have values of λ_i and ϕ_i that are in the neighbourhood of 0.001 and 0.999, respectively. This includes allowing for heterogeneous parameter values between the zealots.

is the number of non-zealot individuals, who are called “civilians” for purposes of clarity in subsequent writing (in the simulations, f_z is always chosen such that $f_z n$ is an integer, with the choice of f_z specified later). For all of this section, mention of opinions will be a reference to the *civilians’ private and expressed opinions*, denoted by the column vectors $\mathbf{y}_c, \hat{\mathbf{y}}_c$. If reference is made to the essentially constant opinions of the zealots, then this will be clearly differentiated; $\mathbf{y}_z(t) \approx \hat{\mathbf{y}}_z(t) \approx \mathbf{1}_{f_z n}$ for all $t \geq 0$ will be used to denote the opinions of the zealots.

Small-world networks approximate real-world networks in terms of transitivity⁵ and the fact that the shortest-path distance between most pairs of vertices is small [Newman, 2010, Chapter 15.1], but do not accurately approximate the degree distribution of many real world networks. On the other hand, scale-free networks are more appropriate to capture the degree distribution of real world networks, but not in terms of transitivity and shortest-path distance. Thus, both small-world and scale-free networks will be studied⁶. The network $\mathcal{G}[\mathbf{W}]$ is generated using standard MATLAB packages for generating small-world and scale-free networks. These packages generate an undirected and unweighted network $\tilde{\mathcal{G}}[\tilde{\mathbf{W}}]$, where $\tilde{w}_{ij} = \tilde{w}_{ji} = 1$ for any $\tilde{w}_{ij} > 0$. For convenience in ensuring the network is strongly connected and that \mathbf{W} is row-stochastic (and thus satisfying Assumption 3.1), the graph $\tilde{\mathcal{G}}[\tilde{\mathbf{W}}]$ is transformed into a directed graph $\mathcal{G}[\mathbf{W}]$ in the following manner. First, a self-loop e_{ii} is added for each $i \in \mathcal{I}$. Then, for every row i , the value of each w_{ij} is drawn randomly from a uniform distribution in the interval $(0, 1)$ if $(v_j, v_i) \in \mathcal{E}$. Once all w_{ij} are determined, the weights w_{ij} are normalised by dividing all entries in row i by $\sum_{j=1}^n w_{ij}$, which ensures that \mathbf{W} is row-stochastic. A consequence of this method is that $w_{ij} > 0 \Leftrightarrow w_{ji} > 0$, which is convenient in facilitating discussion of the results using network statistics such as degree distribution (in this Section 4.3, the degree of individual i is therefore the cardinality of \mathcal{N}_i)⁷.

The $n - f_z n$ civilians are assigned values of $y_i(0)$, ϕ_i and λ_i , picked from beta distributions, which have two parameters α and β . For $\alpha, \beta > 1$, a beta distribution of the variable x is unimodal and satisfies $x \in (0, 1)$, which in turn satisfies Assumption 3.1 regarding λ_i, ϕ_i . The beta distributions for $y_i(0)$, ϕ_i and λ_i are given by $\alpha = 2$, $\beta = 15$; an example distribution of ϕ_i is shown in Fig. 4.8. This means that on average, civilians have low resilience and susceptibility, and initial private opinions close to $y_i(0) = 0.1$. This may represent a society with established and traditional opinions, leading to $\lambda_i \approx 0.1$. However, the topic may be politically sensitive, leading to a low resilience, $\phi_i \approx 0.1$. One example may be the 2017 Australian Postal Survey on the legalisation of same-sex marriage [ABS].

Because the network is large, most simulations will consider the local public opinion model, Eq. (3.4). However, as explained in Chapter 3, the author believes use

⁵For details on transitivity and shortest-path distance, see [Newman, 2010, Chapter 7].

⁶It is standard practise in network science to consider several different graph models, e.g. small-world, scale-free, Erdős-Rényi. A number of standard software packages exist for graph generation, and graphical properties such as transitivity, clustering coefficient etc. are well studied. They are widely accepted, despite each model’s shortcomings (which are different between the models) in resembling real-world networks.

⁷Such a condition on the edge weights is not needed for the theoretical results reported in Chapter 3.

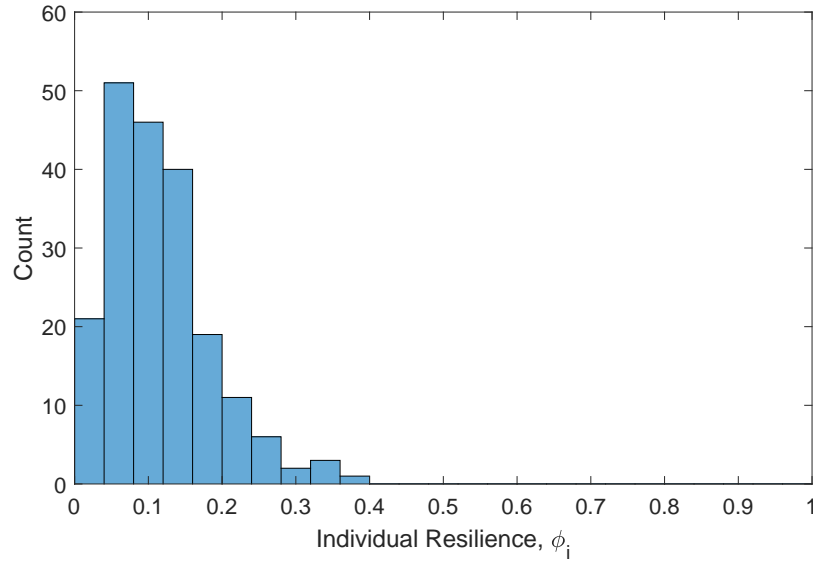


Figure 4.8: Histogram example of distribution of civilian resilience, ϕ_i following a beta distribution with $\alpha = 2$, $\beta = 15$.

of the global public opinion model Eq. (3.2) is also reasonable even though $\hat{y}_{\text{avg}}(t)$ is a global quantity, and in general the network $\mathcal{G}[\mathbf{W}]$ is not all-to-all. The quantity $\hat{y}_{\text{avg}}(t)$ may be available via opinion polls, news and media reports, or in the modern setting, via trends on social media such as hashtags. It is of interest whether availability of the global public opinion has any impact on the opinion evolutions, and this is investigated in Section 4.3.3.

4.3.2 Simulation Set-Up and Results For Small-World Networks

A small-world network is generated as follows. First, a k -regular ring lattice graph of $n = 200$ nodes is generated, with each node connected to k of its nearest nodes, k being an even number. Then, for every edge in the ring lattice graph, an edge is added with probability p between two nodes selected uniformly at random [Newman, 2010]⁸. The small-world networks are generated with $k = 6$ and wiring probability $p = 0.3$. The *mean* degree distribution over all small-world networks of a given type, e.g. $k = 6$ and $p = 0.3$, can be computed analytically, [Newman, 2010, Chapter 15.1.1], and this is shown in Fig. 4.9. For each given network, $f_z n$ individuals are randomly selected to be zealots (examples of appropriate values for f_z are given below). An example simulation, with $f_z = 0.1$ and $n = 200$, is provided in Figs. 4.10

⁸Note that this is slightly different to the original small-world model, which removed an edge and then added an edge between two nodes as describe above, with probability p . Both the network generation method used in this paper, and the original method are described in detail in [Newman, 2010]. The modified generation method ensures that there is no risk that removal of an edge results in disconnected sub-graphs.

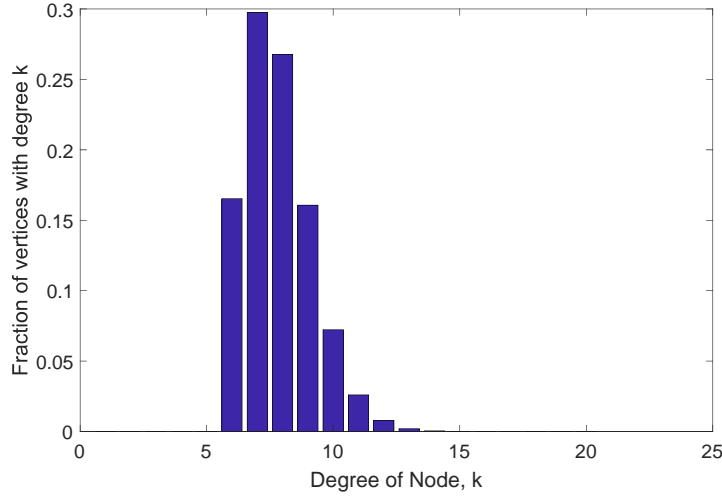


Figure 4.9: Mean of the degree distributions for the small-world networks in the simulations.

and 4.11, showing the evolution of the private and expressed opinions $y(t), \hat{y}(t)$ over time; the red line at $y = 1$ represents the zealots' opinions.

For a given f_z , whose values are given below, 100 small-world networks are generated, with a different set of $W, y(0), \phi, \lambda$ for each network. (See Section 4.3.1 on how these parameters are generated). For these 100 examples, the opinion evolution model Eq. (3.1) and Eq. (3.4) is employed until $t = 1500$ so that the opinions will have converged to the steady state values y^* and \hat{y}^* . The results for $f_z = 0, 0.005, 0.025, 0.05, 0.1$ are shown in Figs. 4.12, 4.13, 4.14, 4.15 and 4.16, respectively. Blue lines represent distributions of final civilian private opinions, y_c^* , and red lines represent distributions of final civilian expressed opinions, \hat{y}_c^* . In each figure, a dotted line represents a single simulation, and specifically, *the distribution of the final civilian opinions* y_c^* and \hat{y}_c^* , normalised so that the sum of the points in the distribution is equal to 1. The zealots' opinions change infinitesimally, and are therefore not recorded or displayed (see Fig. 4.10). The thick solid lines denote the *mean of the distributions of opinions* for the 100 example simulations. It can be seen that while there is significant “fuzziness” in some figures, in the sense that there is variation in the distributions between individual simulations due to the network and associated parameter generation process being random, it is also clear that almost all simulation example distributions follow a shape similar to the mean of the distributions.

4.3.3 Simulation Set-Up and Results For Scale-Free Networks

The same basic set-up as discussed above is used, except now with scale-free networks. These networks were generated using a preferential attachment algorithm [Newman, 2010], with $n = 200$ final desired nodes and $m = 2$ edges attached per

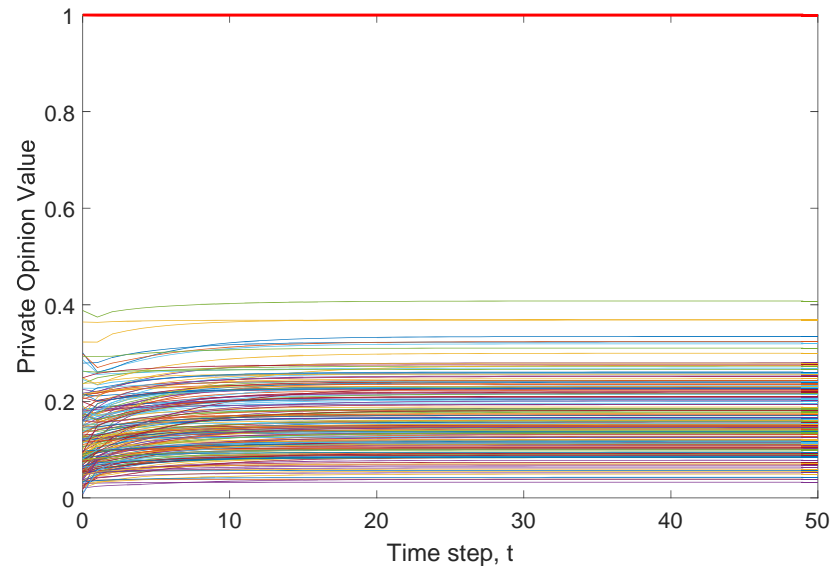


Figure 4.10: Evolution of private opinions, $y_i(t)$, over time t , for an example small-world network with $n = 200$ and 20 zealots. The red bold line at $y = 1$ represents the zealots' opinions. The other coloured lines represent civilian opinions.

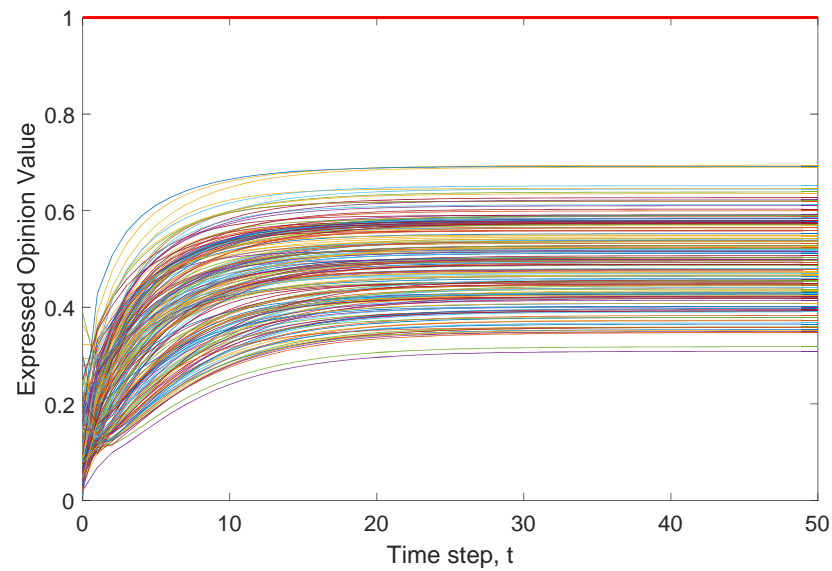


Figure 4.11: Evolution of expressed opinions, $\hat{y}_i(t)$, over time t , for an example small-world network with $n = 200$ and 20 zealots. The red bold line at $y = 1$ represents the zealots' opinions. The other coloured lines represent civilian opinions.

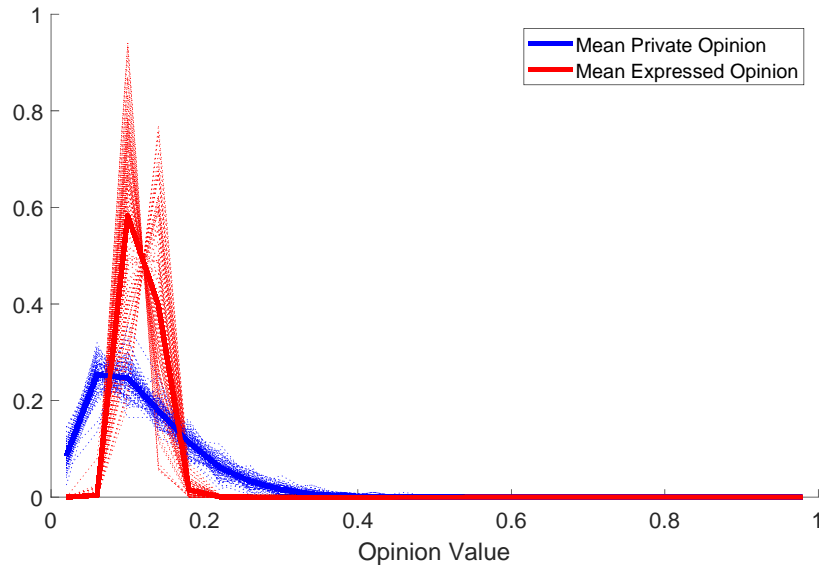


Figure 4.12: Distribution of **civilians'** final private and expressed opinions, y_c^* and \hat{y}_c^* , with 0 zealots in $n = 200$ small-world network.

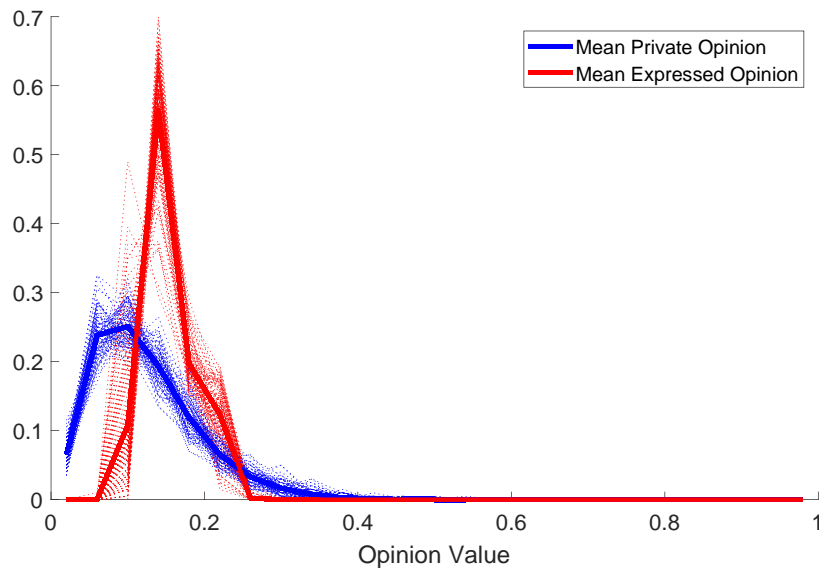


Figure 4.13: Distribution of **civilians'** final private and expressed opinions, y_c^* and \hat{y}_c^* , with 1 zealot in $n = 200$ small-world network.

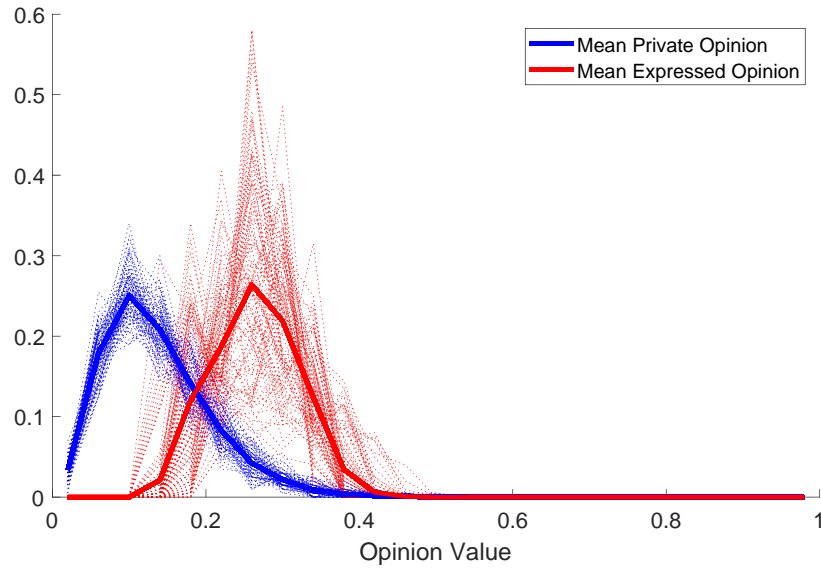


Figure 4.14: Distribution of **civilians'** final private and expressed opinions, y_c^* and \hat{y}_c^* , with 5 zealots in $n = 200$ small-world network.

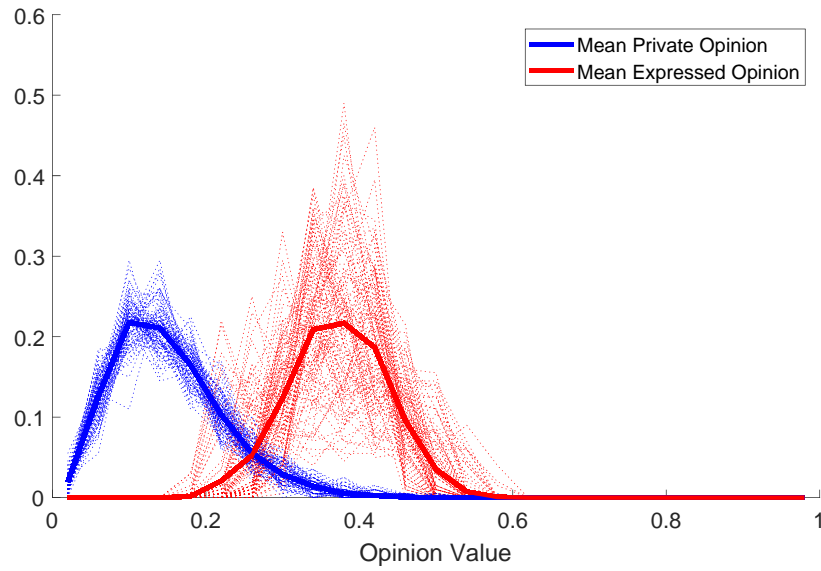


Figure 4.15: Distribution of **civilians'** final private and expressed opinions, y_c^* and \hat{y}_c^* , with 10 zealots in $n = 200$ small-world network.

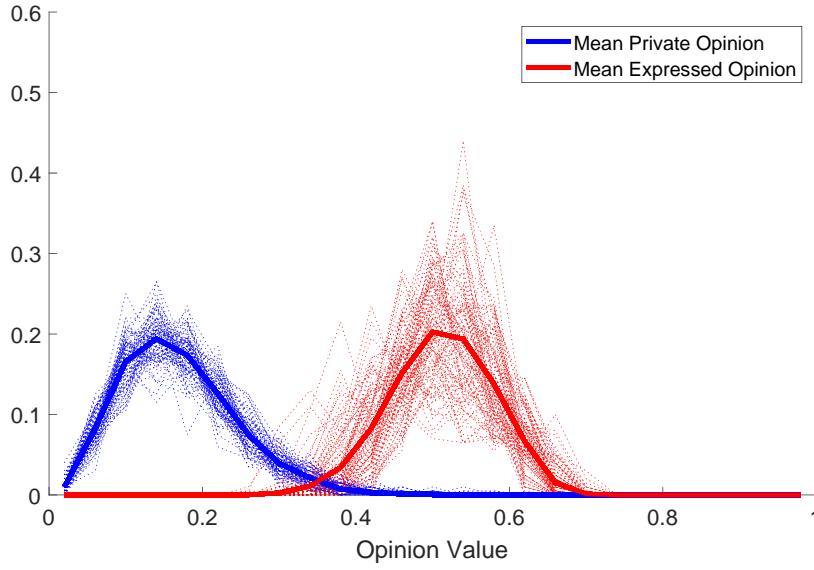


Figure 4.16: Distribution of **civilians'** final private and expressed opinions, y_c^* and \hat{y}_c^* , with 20 zealots in $n = 200$ small-world network.

step of the algorithm⁹. An example of the degree distribution of a scale-free network used in the simulations is given in Fig. 4.17. Under the scheme *Targeted Zealot Placement*, the $f_z n$ highest degree nodes are assigned to be the zealots. Under the scheme *Random Zealot Placement*, $f_z n$ nodes are assigned to be zealots randomly. Intuitively, one would expect that with targeted zealot placement, a greater shift of opinions of civilians would result.

This is borne out by the simulation results, which are now described for both schemes. The edge weights w_{ij} , and civilian initial opinions, resilience, and stubbornness are generated as described in Section 4.3.1. Simulations were run for 100 scale-free networks with $f_z = 0.025$, i.e. 5 zealots in an $n = 200$ individual network, for both zealot placement schemes. This is repeated for both the local public opinion variation, Eq. (3.4), and the global public opinion variation, Eq. (3.2). The results are shown in Fig. 4.21, 4.22, and 4.23. As above, the blue lines represent distributions of final civilian private opinions, y_c^* , and red lines represent distributions of final civilian expressed opinions, \hat{y}_c^* . A dotted line represents a single simulation, normalised so that the sum of the points in the distribution is equal to 1. The thick solid lines denote the *mean of the distributions of opinions* for the 100 example simulations. The global public opinion simulation results using random zealot placement are omitted because there was little difference to the results using the local public opinion update Eq. (3.4) with random zealot placement (see Fig. 4.23). Fig. 4.18, 4.19 and 4.20 show

⁹Roughly speaking, the algorithm begins with a node v_1 . Then, new nodes are added to the network one at a time and edges are added. Specifically, each new node v_i is connected to m nodes v_{j_1}, \dots, v_{j_m} , with $j_1, \dots, j_m < i$. The probability p_j of connecting v_i to v_j is proportional to the degree of v_j prior to any connection of v_i and any nodes added after v_i .

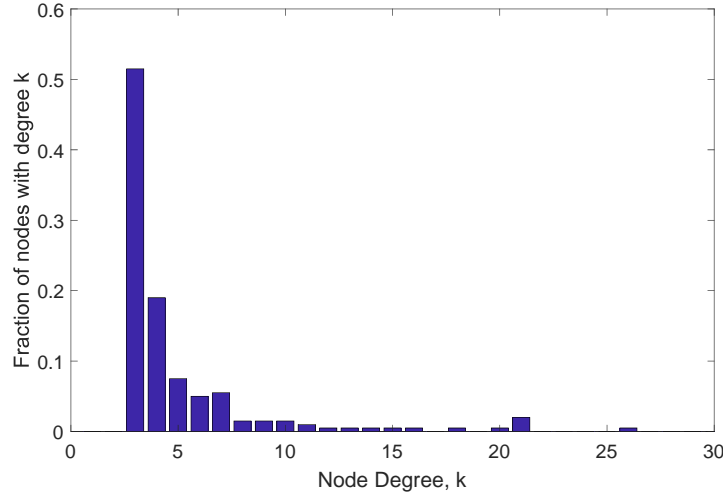


Figure 4.17: Example degree distribution of a scale-free network used in the simulations.

trajectories of the opinions over time for a single simulation. Because civilians are highly stubborn, the private opinions when using local public opinion update and global public opinion update do not differ greatly, and therefore only the private opinions under local public opinion update are shown in Fig. 4.18. On the other hand, there is significant difference in the expressed opinions depending on what update rule is used, as shown in Fig. 4.19 and Fig. 4.20. Discussion of the results will be presented shortly.

4.3.4 Discussion for Small-World Networks

It is clear from Figs. 4.12, 4.13, 4.14, 4.15 and 4.16 that an increase in the number of zealots results in the distributions of y_c and \hat{y}_c diverging. For convenience, denote \mathcal{Z} as the set of zealot nodes of the graph \mathcal{G} . Denote the set of civilians who are 1-hop neighbours of a zealot by \mathcal{C}_1 , i.e. $\mathcal{C}_1 = \{v_j : v_k \in \mathcal{Z}, v_k \in \mathcal{N}_j, j \in \mathcal{V} \setminus \mathcal{Z}\}$. Similarly, for $i = 2, 3, \dots$, denote with the set \mathcal{C}_i the civilian neighbours that are i -hop neighbours of a zealot, i.e. $\mathcal{C}_i = \{v_j : v_k \in \mathcal{C}_{i-1}, v_k \in \mathcal{N}_j, j \in \mathcal{V} \setminus \mathcal{Z}\}$.

First, consider Fig. 4.10 and 4.11 in conjunction with Fig. 4.16. From Fig. 4.10, one can see that the civilians' private opinions do not change by much because the civilians are on average closed to influence, i.e. with small λ_i . This can be verified by looking at Eq. (3.1), where $(1 - \lambda_i)y_i(0)$ is the dominating term on the right hand side, for any civilian i , and for all t . However, civilians on average have low resilience, and are therefore more likely to exhibit *preference falsification* towards their local public opinion. For any civilian $v_j \in \mathcal{C}_1$, his/her local public opinion will be closer to 1 due to the presence of the zealot as a neighbour, and v_j will therefore express an opinion $\hat{y}_j(t)$ closer to 1. Because the zealots' opinions are constant, after each time step, $v_j \in \mathcal{C}_1$ will express an opinion $\hat{y}_j(t)$ that progressively moves ever closer to 1, until equilibrium is reached for v_j . Likewise, civilian $v_k \in \mathcal{C}_2$ observes that his/her

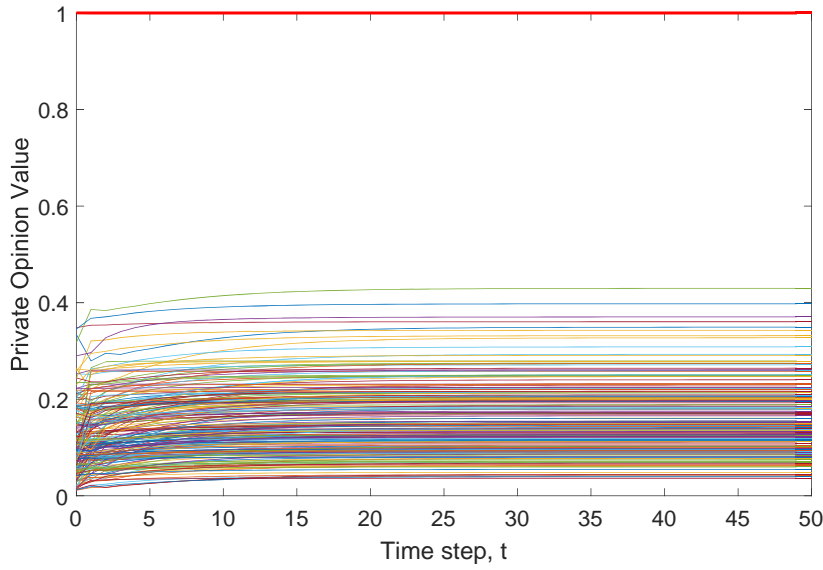


Figure 4.18: Evolution of private opinions, $y_i(t)$, over time t , for an example scale-free network with $n = 200$ and 5 zealots, using **targeted zealot placement**, and **local public opinion**. Notice the similarity to Fig. 4.10 because civilians typically have low susceptibility λ_i . The red bold line at $y = 1$ represents the zealots' opinions. The other coloured lines represent civilian opinions.

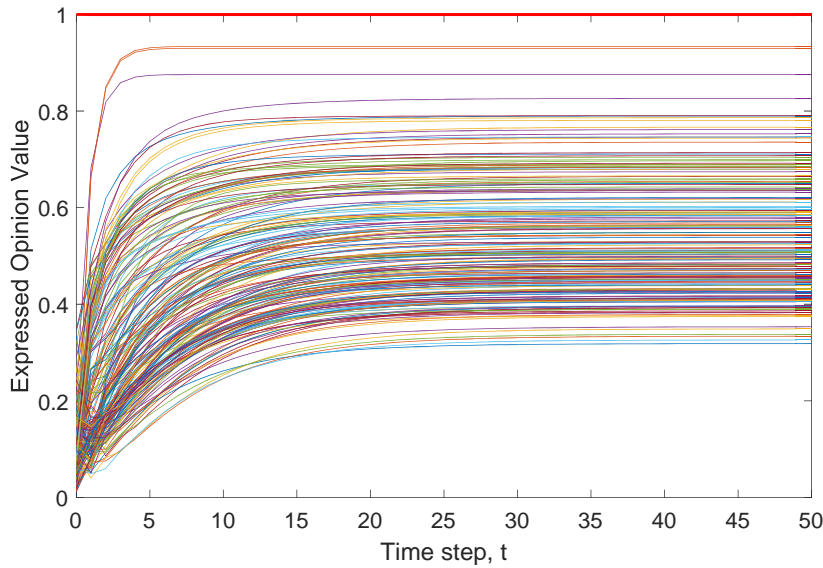


Figure 4.19: Evolution of expressed opinions, $\hat{y}_i(t)$, over time t , for an example scale-free network with $n = 200$ and 5 zealots, using **targeted zealot placement**, and **local public opinion**. The red bold line at $y = 1$ represents the zealots' opinions. The other coloured lines represent civilian opinions.

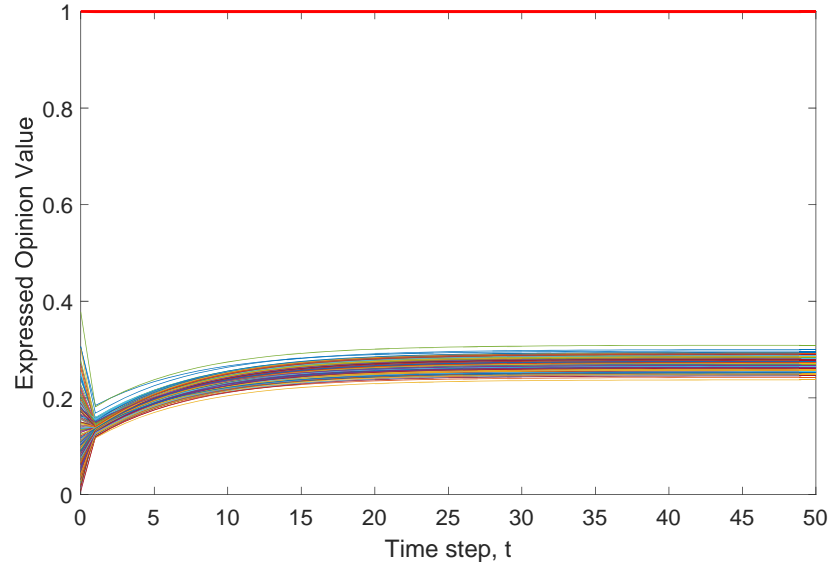


Figure 4.20: Evolution of expressed opinions, $\hat{y}_i(t)$, over time t , for an example scale-free network with $n = 200$ and 5 zealots using **targeted zealot placement**, and **global public opinion**. The red bold line at $y = 1$ represents the zealots' opinions. The other coloured lines represent civilian opinions.

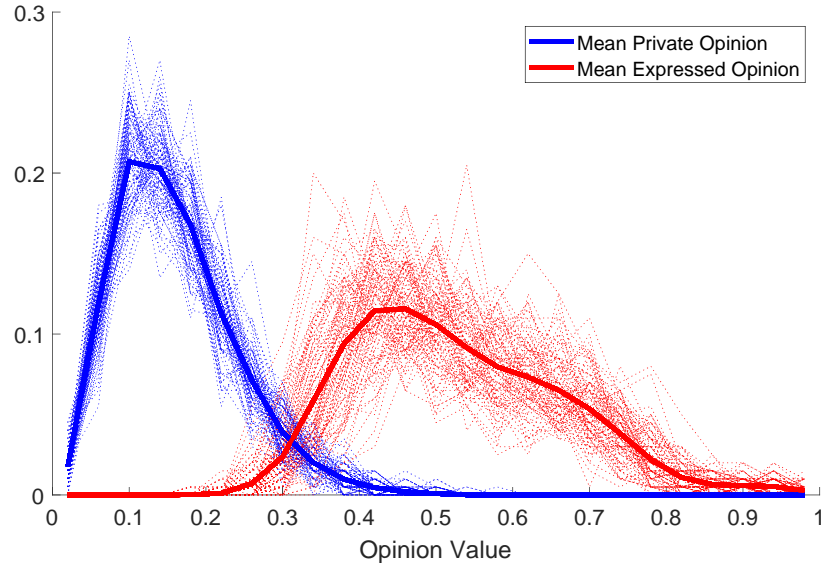


Figure 4.21: Distribution of **civilians'** final private and expressed opinions, y_c^* and \hat{y}_c^* , with 5 zealots in $n = 200$ scale-free network, using **targeted zealot placement** scheme and **local public opinion**.

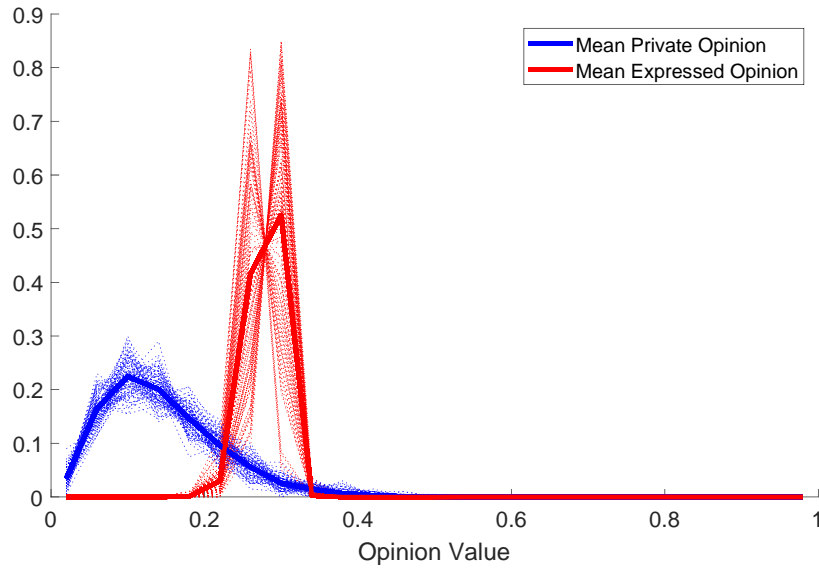


Figure 4.22: Distribution of **civilians'** final private and expressed opinions, y_c^* and \hat{y}_c^* , with 5 zealots in $n = 200$ scale-free network, using **targeted zealot placement** scheme and **global public opinion**.

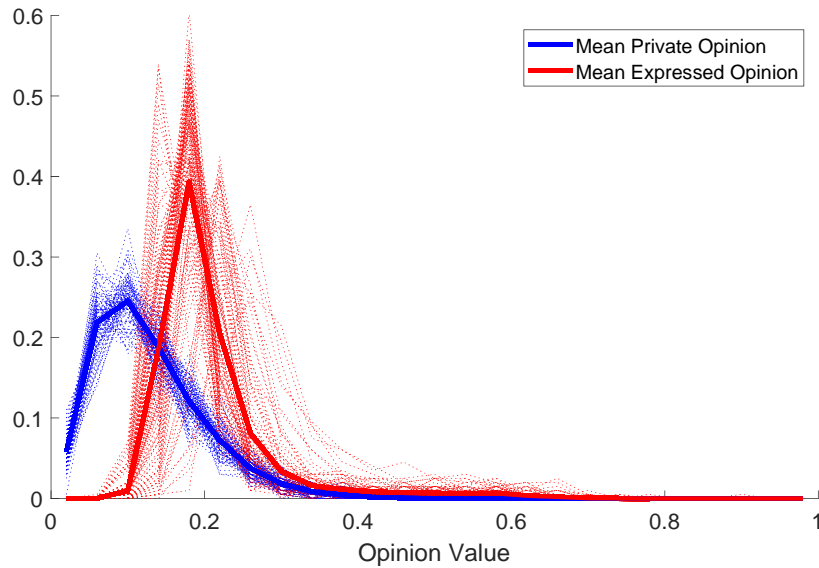


Figure 4.23: Distribution of **civilians'** final private and expressed opinions, y_c^* and \hat{y}_c^* , with 5 zealots in $n = 200$ scale-free network, using **random zealot placement** scheme and **local public opinion**.

neighbour $j \in \mathcal{C}_1$ has expressed an opinion \hat{y}_j that is moving towards 1. This affects the local public opinion of v_k , and in turn, v_k also begins to express an opinion closer to 1. The same holds for $\mathcal{C}_i, i = 3, 4, \dots$ but unsurprisingly, the effects are lessened significantly as the hop distance to a zealot increases. To summarise, at steady state, the civilians are falsifying their expressed opinions y_c^* in the direction of the zealots' opinions because the presence of zealots results in a pressure to conform to a local public opinion which is *not reflective of the true average civilian expressed opinion* (this is especially true for those who are 1- and 2-hop neighbours of the zealots).

Next, notice in Fig. 4.16 that, averaged over 100 simulations, the majority of the civilians hold a final private (true) opinion $y_{i,c}^* \approx 0.15$. Consider a civilian in the network: whether he/she is able to obtain the global public opinion, or just his/her local public opinion, the civilian would see most final expressed opinions at $\hat{y}_{i,c}^* \approx 0.65$. This implies that, all civilians will show strong signs of pluralistic ignorance because they are led to believe that the majority position is approximately 0.65 whereas the true majority position is significantly different, at 0.15. In contrast, Fig. 4.12 shows that when there are no zealots, averaged over 100 simulations, the civilians falsify their final expressed opinions to a negligible extent, but towards the true average civilian public opinion $\hat{y}_{c,avg}$. There is no pluralistic ignorance in this case.

Having established that the presence of zealots can result in pluralistic ignorance and preference falsification towards the zealots' opinion, one turns to study the effect of varying f_z , the fraction of zealots in the network. From Figs. 4.12, 4.13, 4.14, 4.15 and 4.16, it can be concluded that the amount and intensity of pluralistic ignorance and preference falsification toward the zealots' opinion, increases as f_z increases. This is expected. More surprisingly, Fig. 4.14 indicates that it takes only 5 zealots in a $n = 200$ size network, i.e. $f_z = 0.025$, to create a major divergence between the civilians' majority private opinion and expressed opinion. With only 10 zealots, i.e. $f_z = 0.05$ there is significant pluralistic ignorance and preference falsification in most of the civilian population, as shown in Fig. 4.15. It is conjectured that this is due to the small-world effect; any zealot is able to indirectly influence (in a manner explained in the second paragraph of this subsection) a large number of the civilians over short path-lengths (roughly speaking, strength of influence decreases as path-length increases).

4.3.5 Discussion for Scale-Free Networks

First, consider Fig. 4.19, which shows the expressed civilian opinions for a simulation with 5 zealots under the targeted placement scheme in an $n = 200$ scale-free network, where each civilian used local public opinion updating. The level of pluralistic ignorance and preference falsification is significantly higher in Fig. 4.19 than what occurs even with 20 zealots in the small-world network; compare with Fig. 4.11. Note that the private opinions are similar for both scale-free and small-world networks, due to civilians having on average high stubbornness (low λ_i), see Fig. 4.10 versus 4.18. Therefore, no further discussion of private opinions will occur in any detail. One concludes that for local public opinion updating, high degree zealot nodes

can drastically increase pluralistic ignorance and preference falsification in scale-free networks, when compared to small-world networks. On the other hand, when using global public opinion updating, high degree zealot nodes are much less effective, see Fig. 4.19 versus 4.20. This is verified in Fig. 4.21 and 4.22.

This distinction arises because in scale-free networks, most civilians connect to only 2-5 other nodes (in the example in Fig. 4.17, this is 80% of all nodes). This means that under the *targeted placement scheme*, it is highly likely that almost any civilian is a small hop distance from a zealot. Thus, under local public opinion updating, zealots will have a high impact in shaping $\hat{y}_{i,\text{avg}}(t)$ for many civilians, at any time t . This means that, zealots are able to spread misinformation about public opinion to most civilians in the network directly, instead of indirectly as in small-world networks. On the other hand, when using global public opinion updating, the zealots contribute only a small amount to $\hat{y}_{\text{avg}}(t)$. Influence on most civilians is then via the network W ; a large number of civilians will update their private opinion using Eq. (3.1), which will include a zealot as $\hat{y}_{j,z}$ in the term $\sum_{j \neq i}^n w_{ij} \hat{y}_j(t)$. Because civilians are stubborn, this influence is small. As Fig. 4.20 shows, the civilians eventually settle on final expressed opinions which are much more representative of the true civilian public opinion of the network, though there is still a shift toward $y = 1$. Moreover, the distribution of final expressed opinions has much smaller spread when using global public opinion compared to local public opinion, because the low resilience means that civilians express values close to \hat{y}_{avg}^* , see Fig. 4.22 versus 4.21. The relevant theoretical result is in Section 3.3.2.

In contrast to the situation for scale-free networks, extensive simulations in small-world networks found no observable differences between local public opinion updating and global public opinion updating for small-world networks with $k = 6$ and $p = 0.3$. Last, compare the result of 5 zealots for small-world networks, shown in Fig. 4.14 with 5 zealots placed using the *random placement scheme* in scale-free networks with local public opinions, i.e. Fig. 4.23. Clearly, there is lower pluralistic ignorance in scale-free networks when zealots are randomly placed than in small-world networks. This is entirely reasonable, because there is a much higher probability that the zealots in the scale-free network have lower degree. One concludes that indirect influence, as measured by transitivity and shortest-path distance, does have a role in creating pluralistic ignorance, but is less important than high degree zealots.

4.3.6 Key Observations and Insights

There are clear differences in the level of pluralistic ignorance that arise from the different network types. The causes are summarised and explained, with the aid of Fig. 4.24.

For scale-free networks with high degree zealots, most civilians have a low degree (few neighbours) and are likely connected to a zealot, as illustrated in Fig. 4.24a. Thus, a high degree zealot in a scale-free network is a major contributor to the civilian's local public opinion when using Eq. (3.4), but is largely drowned out in the global public opinion; compare Fig. 4.24a and 4.24b. Pluralistic ignorance is greatly

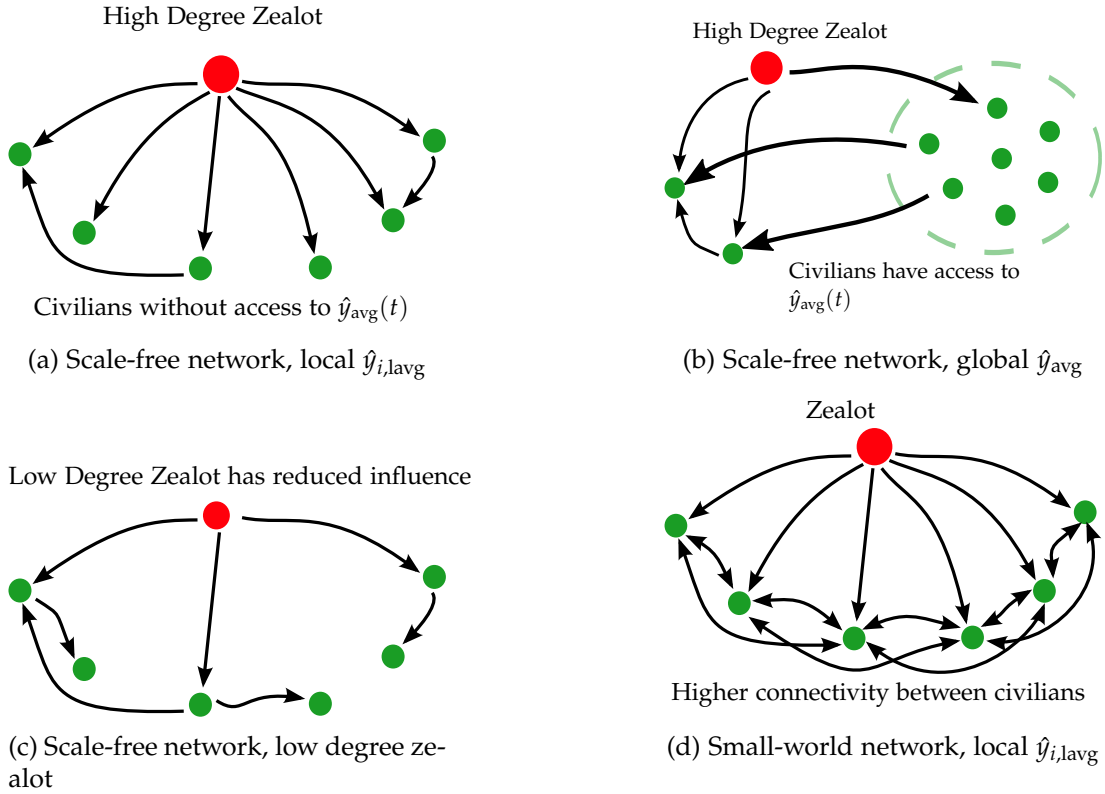


Figure 4.24: The key ways a zealot is able to influence the civilian population for different types of networks are illustrated here. Note that while the edges in the simulations are bidirectional, for clarity, only some key unidirectional edges are shown here.

reduced when civilians have access to the global public opinion (compare Fig. 4.19 and 4.20), with higher levels of agreement in the final expressed opinions. When 5 zealots are placed randomly on a scale-free network the level of pluralistic ignorance is lower than with high degree zealots, regardless of whether Eq. (3.2) or Eq. (3.4) is used. This is because there is a high probability that the zealot has low degree and can directly influence far fewer civilians, see Fig. 4.24c. See also Fig. 4.21, 4.22, 4.23.

In small-world networks, the impact of the zealots is reduced (but not eliminated), compare Fig. 4.12–4.16 to Fig. 4.21–4.23. This is because, as illustrated in Fig. 4.24d, most civilians are connected to several other civilians, so that the effect of the zealot on the local public opinion is diluted. Roughly speaking, it takes 20 zealots in an $n = 200$ small-world network to achieve a similar level of pluralistic ignorance as 5 zealots in scale-free networks; compare Fig. 4.10 and 4.11 to Fig. 4.18 and 4.19. The author conjectures that targeted zealot placement in small-world networks would increase pluralistic ignorance, due to the higher degree of the zealots, but would still be less effective than targeted zealot placement in scale-free networks because of the aforementioned diluting effect.

To summarise, for local public opinion updating, the highest degree nodes are most important in the EPO model because of their direct influence, whereas high

transitivity and shortest-path distance are less important because these determine indirect influence. *Scale-free networks* for which individuals only know local public opinion $\hat{y}_{i,\text{avg}}$ are particularly susceptible to disruption by zealots, whereas *small-world networks* are more robust. Availability of global public opinion $\hat{y}_{\text{avg}}(t)$ greatly reduces the level of pluralistic ignorance.

Connecting civilians together, even if they have differing opinions, and ensuring reliable access to global $\hat{y}_{\text{avg}}(t)$, are therefore good ways to counter high degree zealots and reduce pluralistic ignorance, even if private $y_i(t)$ are not known. This is particularly important in the age of digital information and social media where individuals increasingly interact only with others of similar opinions and consume media from partisan sources. The author notes that misinformation has been identified in [Kuran, 1997] as a key reason for pluralistic ignorance and preference falsification. From one point of view, one might say that the civilians have incorrect knowledge about the public opinion; the true public opinion is much closer to 0, but the zealots are distorting this information. In complex network theory, high degree nodes play important roles in determining the behaviour of many dynamical processes [Newman, 2010]. The significant influence of zealots spreading falsehoods in mass consumed media has been identified [McCright and Dunlap, 2017]. In addition, the “minority effect” has been reported, whereby a small but vocal minority are able to exert influence over a larger community [Smaldino and Epstein, 2015; Wood et al., 1994; Yildiz et al., 2013]. Here, a mathematical, agent-based examination and explanation of these ideas in the context of opinion formation in influence networks has been provided.

4.4 Conclusions

This chapter has reviewed two established results from social psychology literature; the Asch experiments and the phenomenon of pluralistic ignorance. Using a combination of theoretical and simulation based analysis, these two issues have been examined using the EPO model. In particular, Section 4.2 showed that the three different types of individuals observed in Asch’s experiments could be accurately captured and predicted by suitable values of stubbornness and resilience for that individual. Also of note is that the model provides an explanation for why Asch observed no proportional increase in the distortion of the individual’s judgment or action when the majority size was increased. Section 4.3 showed how zealots could generate massive pluralistic ignorance in a general population of individuals. Through extensive simulation and analysis, differences between (i) scale-free and small-world networks, and (ii) global public opinion updating and local public opinion updating were identified. As a result, the author concludes that increasing connectivity among civilians and making the global public opinion more readily available are good ways to reduce the influence of zealots. A number of possible future works are detailed in Chapter 10.

Part II

Evolution of Individual Social Power

Evolution of Social Power in Networks with Constant Topology

Part Summary

Part II moves away from the EPO model, and turns to study of the recently proposed DeGroot–Friedkin model [Jia et al., 2015], which describes how an individual’s self-weight (referred to as social power) evolves as a network discusses opinions on a sequence of topics (in contrast, the EPO model considers discussion of a single topic). Chapters 5 and 6 study convergence properties of the model with constant and dynamic network topology, respectively. Chapter 7 studies a network modification problem on star graphs, which are special networks where a single individual eventually holds all of the social power. This part is concluded with Chapter 8, which applies a recent nonlinear mapping convergence result to the DeGroot–Friedkin model.

Chapter Summary

This chapter considers the DeGroot–Friedkin model on networks with a constant topology, an assumption which is relaxed in the following chapter. A novel analysis framework is built using nonlinear contraction theory, which allows the drawing of a new and general exponential convergence result. Specifically, it will be shown that each individual’s social power converges exponentially fast to a constant value that depends only on the network topology. Previous results establish asymptotic, but not exponential convergence. Additional analysis yields an explicit upper bound on the individual’s social power at equilibrium that is dependent on the graph topology, and the convergence rates for a general class of graph topologies are obtained.

5.1 Introduction

In the previous chapters, focus has been placed on the study of how the opinions of individuals evolve via interpersonal influence. In each instance (from the DeGroot

and Friedkin–Johnsen models to the EPO model), the influence weights w_{ij} have been assumed to be time-invariant. However, this time-invariant property clearly does not hold over long periods of time. Many models assume that the weights w_{ij} evolve according to certain social processes that often depend on the opinion values, e.g. the Hegselmann–Krause model¹ [Hegselmann and Krause, 2002] assumes that the w_{ij} are dependent on the magnitude of $|y_i - y_j|$. Part II now turns to a model which examines how w_{ii} , i.e. an individual’s self-weight (or social power) evolves over a sequence of topic discussions via the social process of *reflected self-appraisal*.

The concept of *social power* or *social influence* has been integral throughout the development of many agent-based models of opinion dynamics. Indeed, French Jr’s seminal paper [French Jr, 1956] was an attempt to quantitatively study an individual’s social power in a group discussion. Broadly speaking, and in the context of opinion dynamics, an individual’s social power is the amount of influence an individual has on the overall opinion discussion. It is expected that the social power of an individual in a group can change over time as group members interact and are influenced by each other. Recently, the DeGroot–Friedkin model was proposed in [Jia et al., 2015] to study the dynamical evolution of an individual’s social power as a social network discusses opinions on a sequence of topics (also known as issues).

The discrete-time DeGroot–Friedkin model is a two-stage model. In the first stage, individuals update their opinions on a particular issue using the DeGroot model (see Section 2.3.1), and in the second stage, each individual’s self-weight (referred to as social power) w_{ii} for the next issue is updated. Under some mild conditions on the network topology, the opinions reach a consensus on every issue. At the end of the discussion of an issue, i.e., when opinions have effectively reached a consensus, each individual undergoes the socio-psychological process of *reflected self-appraisal* (as detailed in the seminal work [Cooley, 1992]) to determine his/her impact or influence on the final consensus value of opinion. Such a mechanism is well accepted as a hypothesis [Shrauger and Schoeneman, 1979; Gecas and Schwalbe, 1983] and has been empirically validated [Yeung and Martin, 2003]. Immediately before discussion on the next issue, each individual i self-appraises and updates individual social power w_{ii} according to the impact or influence i had on the discussion of the previous issue. In updating w_{ii} , individual i also updates the weight i accords to i ’s neighbours’ opinions w_{ij} , by scaling using a *relative interaction matrix*. This is to ensure that the influence matrix of the DeGroot process, W , remains row-stochastic for the next issue. The relative interaction matrix encodes the relative “trust” or “friendship” between the individuals in the network, and in the original model [Jia et al., 2015], the relative interaction matrix was assumed to be constant over all issues, and constant throughout the opinion discussion on any given issue. The primary objective of the DeGroot–Friedkin model is to *study the dynamical evolution of the individual social powers over the sequence of issues, with self-appraisal occurring after each discussion*.

In order to fully capture the novel contributions of this chapter, the DeGroot–Friedkin model must be formally introduced. This is done immediately below, along

¹The reader is referred to Chapter 1 for a review of some relevant models.

with a review of existing results. Note that Chapters 5 through 8 in Part II use the terms “self-weight”, “social power”, “self-confidence” and “individual social power” interchangeably, and the terms “topic” and “issue” interchangeably.

5.1.1 The DeGroot–Friedkin Model

Consider a social network which discusses a sequence of issues, with the set of indices for the sequential issues denoted by $\mathcal{S} = \{0, 1, 2, \dots\}$. For a given issue $s \in \mathcal{S}$, the opinions are assumed to evolve according to the DeGroot model (with constant influence weights throughout the discussion of the issue). At the end of the discussion of a particular issue s (i.e. when the DeGroot model has effectively reached steady state), each individual undergoes reflected self-appraisal, with “reflection” referring to the fact that self-appraisal occurs following the completion of discussion on the particular issue s . Each individual then updates his/her own self-weight, and discussion begins on the next issue $s + 1$ (using the DeGroot model but now with adjusted influence weights).

The structure of the social network is captured by the directed graph $\mathcal{G}[C] = (\mathcal{V}, \mathcal{E}[C], C)$, where $C \in \mathbb{R}^{n \times n}$ is the “relative interaction matrix” (the relation of C to the influence matrix W in the DeGroot model is explained below). The matrix C has nonnegative entries c_{ij} , termed “relative interpersonal weights” in [Jia et al., 2015]. It is assumed that $c_{ii} = 0$ (i.e., there are no self-loops), and the restriction that $\sum_{j \in \mathcal{N}_i} c_{ij} = 1$ is imposed (i.e., C is a row-stochastic matrix). The word “relative” is used to reflect that c_{ij} can be considered as a percentage of the total weight or trust individual i places on individual j compared to all of individual i ’s neighbours. A mild assumption is now made on the relative interpersonal interaction network $\mathcal{G}[C]$, which will hold not only for this chapter but all other chapters in Part II.

Assumption 5.1. *The graph $\mathcal{G}[C]$, with $n \geq 3$ nodes, is strongly connected and no node v_i has a self-loop. Furthermore, the relative interaction matrix C is row-stochastic.*

Because there are no self-loops, C has zero diagonal entries. The strongly connected nature of $\mathcal{G}[C]$ implies that C is irreducible. The *dominant left eigenvector* of the row-stochastic C is denoted by $\gamma^\top > 0$ satisfying $\gamma^\top \mathbf{1}_n = 1$ (see Section 2.2 in Chapter 2). It should be noted that this assumption was in place in [Jia et al., 2015] by and large throughout the development of the DeGroot–Friedkin model, though a recent work has extended the result to consider graphs with reducible C [Jia et al., 2017b]. A key portion of the works of Part II focus on using new techniques to extend convergence results for the model (including for dynamically changing topology), for which the strongly connected assumption is appropriate.

It is suggested that the reader review the definition of the n -dimensional simplex Δ_n , and related definitions for $\tilde{\Delta}_n$ and $\text{int}(\Delta_n)$ in Section 2.1, as these will appear repeatedly over the next four chapters. Next, the mathematical modelling of the opinion dynamics for an issue and the updating of self-weights from one issue to the next, are explained.

5.1.1.1 DeGroot Consensus of Opinions

For each issue $s \in \mathcal{S}$, individual i 's opinion $y_i(s, t) \in \mathbb{R}$ evolves according to the DeGroot process (see Section 2.3.1)

$$y_i(s, t+1) = w_{ii}(s)y_i(s, t) + \sum_{j \neq i}^n w_{ij}(s)y_j(s, t).$$

Note that $\forall i, j$, $w_{ij}(s) \in [0, 1]$ is constant for any given s , but changes between each s due to the reflected self-appraisal process (as explained below). The parameter $w_{ii}(s)$ is the “self-weight” or “social power” of individual i . The opinion dynamics for the entire social network can be expressed as

$$\mathbf{y}(s, t+1) = \mathbf{W}(s)\mathbf{y}(s, t), \quad (5.1)$$

where $\mathbf{y}(s, t) = [y_1(s, t), \dots, y_n(s, t)]^\top$ is the vector of opinions of the n individuals for issue s , at time instant t . The original DeGroot model arises simply when $\mathcal{S} = \{0\}$ (i.e., only one issue is discussed).

Let the self-weight (individual social power) of individual i be denoted by $x_i(s) \triangleq w_{ii}(s) \in [0, 1]$ (the i^{th} diagonal entry of $\mathbf{W}(s)$) [Jia et al., 2015], with the social power vector of the individuals in the network given as $\mathbf{x}(s) = [x_1, \dots, x_n]^\top$. For a given issue s , the influence matrix $\mathbf{W}(s)$ is defined as

$$\mathbf{W}(s) = \mathbf{X}(s) + (\mathbf{I}_n - \mathbf{X}(s))\mathbf{C}, \quad (5.2)$$

where \mathbf{C} is the relative interaction matrix of the social network $\mathcal{G}[\mathbf{C}]$, and the matrix $\mathbf{X}(s) \triangleq \text{diag}(x_i(s))$. From the fact that \mathbf{C} is row-stochastic with zero diagonal entries, Eq. (5.2) implies that $\mathbf{W}(s)$ is a row-stochastic matrix. An assumption² is imposed that the starting self-weights $x_i(0) \geq 0$ satisfy $\sum_{i=1}^n x_i(0) = 1$, i.e. $\mathbf{x}(0) \in \Delta_n$. It has been shown in [Jia et al., 2015] that under this initial condition assumption and Assumption 5.1, $\mathbf{W}(s)$ defined as in Eq. (5.2) has the property that for any given $s \in \mathcal{S}$, there holds $\lim_{t \rightarrow \infty} \mathbf{y}(s, t) = (\boldsymbol{\zeta}(s)^\top \mathbf{y}(s, 0))\mathbf{1}_n$. Here, $\boldsymbol{\zeta}(s)^\top$ is the unique nonnegative left eigenvector of $\mathbf{W}(s)$ associated with the eigenvalue 1, normalised such that $\mathbf{1}_n^\top \boldsymbol{\zeta}(s) = 1$. That is, for every issue, the opinions converge to a constant consensus value (see Chapter 2). Next, the model for the updating of $\mathbf{W}(s)$ (specifically $w_{ii}(s)$) via a reflected self-appraisal mechanism) is introduced.

5.1.1.2 Friedkin's Self-Appraisal Model for Determining Self-Weight

The Friedkin component of the model proposes that at the end of the discussion of issue s , the self-weight vector updates as

$$\mathbf{x}(s+1) = \boldsymbol{\zeta}(s), \quad (5.3)$$

²This assumption is not strictly required, and will be relaxed in Chapter 6.

where $\zeta(s)$ was defined below Eq. (5.2). Note that $\zeta(s)^\top \mathbf{1}_n = 1$ implies that $\mathbf{x}(s) \in \Delta_n$, i.e., $\sum_{i=1}^n x_i(s) = 1$ for all s . From Eq. (5.2), and because \mathbf{C} is row-stochastic, it is apparent that by adjusting $w_{ii}(s+1) = \zeta_i(s)$, individual i also scales $w_{ij}(s+1)$, $j \neq i$ using c_{ij} to be $(1 - w_{ii}(s+1))c_{ij}$ to ensure that $\mathbf{W}(s)$ remains row-stochastic. It should now become apparent to the reader why the c_{ij} are referred to as “relative interaction weights”; as individual i ’s self-weight changes over the sequence of issues \mathcal{S} , i uses the relative interaction weight c_{ij} to scale the influence weight $w_{ij}(s)$ accorded to neighbour j . The objective of the DeGroot–Friedkin model can therefore be summarised as follows

Objective 5.1. Consider a network $\mathcal{G}[\mathbf{C}]$ of n individuals that discuss a sequence of issues indexed by $\mathcal{S} = \{0, 1, 2, \dots\}$, using the DeGroot process Eq. (5.1) for each issue s , with the influence matrix $\mathbf{W}(s)$ determined according to Eq. (5.2). Suppose that at the end of issue $s \in \mathcal{S}$, the vector of social power $\mathbf{x}(s)$ defined above Eq. (5.2) changes according to Eq. (5.3). Then, the objective is to investigate the dynamical evolution of $\mathbf{x}(s)$ over the sequence of issues \mathcal{S} , and determine what (if any) parameters of $\mathcal{G}[\mathbf{C}]$ determine the evolution of $\mathbf{x}(s)$.

Remark 5.1 (Social Power). The precise motivation behind using Eq. (5.3) to update $\mathbf{x}(s)$ is heavily explored in [Jia et al., 2015], but a brief overview is provided here in the interest of making Part II self-contained. As discussed in Section 5.1.1.1, for any given s , there holds $\lim_{t \rightarrow \infty} \mathbf{y}(s, t) = (\zeta(s)^\top \mathbf{y}(s, 0)) \mathbf{1}_n$. In other words, for any given issue s , the opinions reach a consensus value $\zeta(s)^\top \mathbf{y}(s, 0)$ equal to a convex combination of the individuals’ initial opinion values $\mathbf{y}(s, 0)$ for that issue. The elements of $\zeta(s)$ are the convex weights. For a given issue s , $\zeta_i(s)$ is therefore a precise manifestation of individual i ’s social power or influence in the social network, as it is a measure of the ability of individual i to determine the outcome of a discussion, relative to other individuals [Cartwright, 1959]. The reflected self-appraisal mechanism in Eq. (5.3) therefore describes an individual (a) observing how much power he/she had on the discussion of issue s , captured by the nonnegative quantity $\zeta_i(s)$, and (b) for the next issue $s+1$, adjusting his/her self-confidence to be equal to this power, i.e. $x_i(s+1) = w_{ii}(s+1) = \zeta_i(s)$. For a constant \mathbf{W} , the entries ζ_i of the dominant left eigenvector ζ^\top is closely related to what is sometimes referred to as the eigenvector centrality of node i , which is a measure of i ’s importance in the network (in general, networks $\mathcal{G}[\mathbf{W}]$ do not have row-stochastic \mathbf{W} and may not be strongly connected, meaning the eigenvector centralities are not precisely equivalent to ζ_i). Eigenvector centrality [Newman, 2010] is of importance for a number of network problems including Google’s PageRank Algorithm [Brin and Page, 1998], and measuring integration in financial markets [Xu and Corbett, 2015]. An extensive survey of centrality in networks is given in [Gleich, 2015].

Remark 5.2. The model is centralised in the sense that individuals are able to observe and detect their impact relative to every other individual in the opinion discussions process (as measured by $\zeta^\top(s)$), which indicates that the DeGroot–Friedkin model is best suited for networks of small or moderate size. Such networks are found in many decision making groups such as boards of directors, government cabinets or jury panels. A distributed model, referred to as the “modified DeGroot–Friedkin model”, was studied in discrete-time in [Xu et al., 2015; Xia et al., 2016b]. In continuous-time, [Chen et al., 2017] studied a model referred to

as the “distributed DeGroot–Friedkin model”. Dynamic topology, but restricted to doubly-stochastic relative interaction matrices, was studied in [Xia et al., 2016b].

Remark 5.3 (Time-scales). *The DeGroot–Friedkin model assumes the opinion dynamics process operates on a different time-scale than that of the reflected appraisal process. This allows for a simplification in the modelling and is reasonable since separate time-scales merely imply that the network reaches a consensus on opinions on one issue before beginning the process of reflected self-appraisal and then moving onto discussion on the next issue. For a single issue, convergence to consensus occurs exponentially fast (or in finite time in some instances [Hendrickx et al., 2015]). The distributed DeGroot–Friedkin model studies the case when the time-scales are comparable, but the analysis of the distributed model is much more involved, and has not yet reached the same level of understanding as for the original model.*

5.1.1.3 Existing Results

It was shown in [Jia et al., 2015, Lemma 2.2] that the system Eq. (5.3) is equivalent to the discrete-time system³

$$\mathbf{x}(s+1) = \mathbf{F}(\mathbf{x}(s)), \quad (5.4)$$

where the nonlinear map $\mathbf{F}(\mathbf{x})$ is defined as

$$\mathbf{F}(\mathbf{x}) = \begin{cases} \mathbf{e}_i & \text{if } \mathbf{x} = \mathbf{e}_i \text{ for any } i \\ \alpha(\mathbf{x}) \begin{bmatrix} \frac{\gamma_1}{1-x_1} \\ \vdots \\ \frac{\gamma_n}{1-x_n} \end{bmatrix} & \text{otherwise} \end{cases}, \quad (5.5)$$

with $\alpha(\mathbf{x}) = 1 / \sum_{i=1}^n \frac{\gamma_i}{1-x_i}$ where $\boldsymbol{\gamma}^\top = [\gamma_1, \dots, \gamma_n]$ is the dominant left eigenvector⁴ of \mathbf{C} . Note that $\sum_i F_i = 1$.

A special topology studied in [Jia et al., 2015] is termed “star topology”, the definition and relevance of which follow.

Definition 5.1 (Star topology). *A strongly connected graph⁵ $\mathcal{G}[\mathbf{C}]$ is said to have star topology if \exists a node v_i , called the centre node, such that every edge of $\mathcal{G}[\mathbf{C}]$ is either to or from v_i .*

³The system Eq. (5.1) with $\mathbf{W}(s)$ defined in Eq. (5.2) updating according to Eq. (5.3) can be viewed as a hybrid system with state-dependent switching, which can be challenging to analyse. It is unclear to the author whether the following contraction analysis in Section 5.2 may be useful for analysis of other hybrid system with state-dependent switching, since the system in this chapter is a special case, with switching occurring once $\mathbf{y}(s, t)$ has reached a steady state.

⁴It is clear that the entries γ_i play a role in the evolution of social power (the precise role is investigated throughout Part II). As detailed in [Jia et al., 2015; Proskurnikov and Tempo, 2017], $\boldsymbol{\gamma}$ is closely related to the PageRank measure for webpage ranking on the World Wide Web. Much of the PageRank literature focuses on its decentralised computation [Ishii and Tempo, 2014; You et al., 2017].

⁵While it is possible to have a star graph that is not strongly connected, the chapters of Part II, similarly to [Jia et al., 2015], deals only with strongly connected graphs.

The irreducibility of C implies that a star \mathcal{G} must include edges in both directions between the centre node v_i and every other node $v_j, j \neq i$. A lemma and a theorem (the key results of [Jia et al., 2015]) are now provided regarding the convergence of $F(\mathbf{x}(s))$ as $s \rightarrow \infty$, along with a fact useful for analysis throughout the paper.

Lemma 5.1 ([Jia et al., 2015, Lemma 3.2]). *Suppose that $n \geq 3$, and suppose further that $\mathcal{G}[C]$ has star topology, which without loss of generality has centre node v_1 . Let $\mathcal{G}[C]$ satisfy Assumption 5.1. Then, for the system Eq. (5.4), there holds $\lim_{s \rightarrow \infty} \mathbf{x}(s) = \mathbf{e}_1, \forall \mathbf{x}(0) \in \tilde{\Delta}_n$.*

This implies that $\forall \mathbf{x}(0) \in \tilde{\Delta}_n$, a network with star topology converges to an “autocratic configuration” where the centre individual 1 holds all of the social power⁶.

Fact 5.1. [Jia et al., 2015, Lemma 2.3] *Suppose that $n \geq 3$ and let γ^\top be the dominant left eigenvector of $C \in \mathbb{R}^{n \times n}$, satisfying Assumption 5.1. If $\mathcal{G}[C]$ is a star topology graph, then $\gamma_i = 0.5$ where i is the centre node and $\gamma_j < 0.5$ for $j \neq i$; otherwise, $\|\gamma\|_\infty < 0.5$.*

Theorem 5.1 ([Jia et al., 2015, Theorem 4.1]). *For $n \geq 3$, consider the DeGroot–Friedkin dynamical system Eq. (5.4) with $\mathcal{G}[C]$ satisfying Assumption 5.1. Assume further that the graph $\mathcal{G}[C]$ does not have star topology. Then,*

- (i) *For all initial conditions $\mathbf{x}(0) \in \tilde{\Delta}_n$, the self-weights $\mathbf{x}(s)$ converge to \mathbf{x}^* as $s \rightarrow \infty$, where $\mathbf{x}^* \in \text{int}(\Delta_n)$ is the unique fixed point satisfying $\mathbf{x}^* = F(\mathbf{x}^*)$.*
- (ii) *There holds $x_i^* < x_j^*$ if and only if $\gamma_i < \gamma_j$ for any i, j , where γ_i is the i^{th} entry of the dominant left eigenvector γ^\top . There holds $x_i^* = x_j^*$ if and only if $\gamma_i = \gamma_j$.*
- (iii) *The unique fixed point \mathbf{x}^* is determined only by γ , and is independent of the initial conditions $\mathbf{x}(0)$.*

Remark 5.4. *Since the DeGroot model was introduced in [DeGroot, 1974], many different opinion dynamics models have been proposed, of increasing sophistication. Some of these were covered in Chapter 1. However, the DeGroot model continues to be of relevance; the recent paper [Becker et al., 2017] applied the DeGroot model to show how discussion over social networks could improve the “wisdom of crowd” effect, then experimentally validated the results. In future, one could replace Eq. (5.1) with other opinion dynamics models. The key difficulty will be in defining the influence of the individuals, currently captured by $\zeta(s)^\top$ (see Remark 5.1) and obtaining a system Eq. (5.4) with an analysable map F . For the Friedkin–Johnsen model, experimental and simulation results [Friedkin et al., 2016a] are available but theoretical study has proved to be extremely difficult, with limited (at least to this point) results [Mirtabatabaei et al., 2014].*

Two useful properties of $F(\mathbf{x})$ established in [Jia et al., 2015] are now presented.

Property 5.1. *The map $F(\mathbf{x})$ in Eq. (5.5) is continuous on Δ_n .*

⁶A second class of special topologies exists, where $\mathcal{G}[C]$ is such that C is doubly-stochastic, i.e. with row and column sums equal to 1. Then, $\gamma = \mathbf{1}_n/n$ and from Theorem 5.1, the social power vector $\mathbf{x}(s)$ converges to $\mathbf{x}^* = \mathbf{1}_n/n$, known in [Jia et al., 2015] as a democratic configuration. However, this thesis does not focus on this class of topologies, though all results in Chapters 5 and 6 do continue to hold.

If $\mathcal{G}[C]$ does not have star topology, then the following *contraction-like* property holds [Jia et al., 2015, pp. 390, Appendix F].

Property 5.2. *For the system Eq. (5.4), there exists a positive constant r such that for all $r' < r$, $x_i(s) \leq 1 - r'$ implies $x_i(s+1) < 1 - r'$, for all i . Defining the set $\mathcal{A} = \{x \in \Delta_n : 1 - r \geq x_i \geq 0, \forall i \in \{1, \dots, n\}\}$, one has $F(\mathcal{A}) \subset \mathcal{A}$.*

This is termed a *contraction-like* property so as not to confuse the reader with the main result established in Section 5.2 using nonlinear contraction analysis; the above property establishes a contraction only near the boundary of the simplex Δ_n , since the proof of this property establishes that r is small. The size of this “contraction-like” region is explicitly obtained in Section 5.3 by obtaining an exact value for r . Last, a third easily verifiable property is presented.

Property 5.3. *For the system Eq. (5.4), if $x(s_1) \in \tilde{\Delta}_n$ for some $s_1 < \infty$, then $x(s) \in \text{int}(\Delta_n)$ for all $s > s_1$.*

Proof. Since $x(s_1) \in \tilde{\Delta}_n$, $\exists j : 0 < x_j(s_1) < 1$. In addition, $\gamma_i > 0, \forall i$ because C is irreducible. It follows that $\alpha(x(s_1)) > 0$, and $x_i(s_1 + 1) > 0, \forall i$. Along with Property 5.2, it follows that $x(s) \in \text{int}(\Delta_n) \forall s > s_1$. \square

This completes the formal introduction of the original DeGroot–Friedkin model, which will be studied (and extended) in Part II of this thesis. In this chapter, a novel analysis method based on nonlinear contraction theory is used in Section 5.2 to conclude that for non-star graphs, convergence to the fixed point x^* is in fact exponentially fast. The work [Jia et al., 2015], i.e. Theorem 5.1, was able to conclude asymptotic, but not exponential, convergence. In addition, the stability of the equilibria e_i are characterised; it is shown that these equilibria (in which individual i holds all of the social power) are unstable except if individual i is the centre of a star graph. In that case, e_i is an asymptotically stable equilibrium, and exponential stability is ruled out. Section 5.3 significantly extends the “contraction-like” Property 5.2. As a result of the extension, an analytic upper bound is obtained on an individual’s social power at equilibrium; the bound for x_i^* depends only on γ_i . Also, the convergence rate for a class of C is determined, which is again linked to γ . These give important insights into how the topology, and in particular C , affect both the transient dynamics and limiting values of each individual’s social power. Importantly, the nonlinear contraction framework also greatly simplifies the analysis for dynamic, or issue-varying $C(s)$, as will become evident in Chapter 6.

5.1.2 Chapter Organization

The rest of this chapter is organized as follows. Section 5.2 uses nonlinear contraction analysis to establish an exponential convergence result for the original DeGroot–Friedkin model. Section 5.3 conducts further analysis on the DeGroot–Friedkin model to obtain an analytic upper bound on an individual’s social power, and draw conclusions on the convergence rates for a class of graph topologies. Concluding remarks are given in Section 5.4

5.2 Exponential Convergence to Constant Social Power

A fundamental convergence result of the system Eq. (5.4) will now be presented. In the original work [Jia et al., 2015], LaSalle's Invariance Principle was used to prove an asymptotic convergence result. Moreover, a Lyapunov function that was dependent on the unique fixed point \mathbf{x}^* of the map F was used. Neither of these techniques can be extended to general dynamic and issue-varying $C(s)$, which is addressed in the next chapter. This chapter strengthens the result by establishing exponential convergence, and builds an analysis framework upon which dynamic $C(s)$ can be studied in the next chapter.

Remark 5.5. *It should be noted that as a consequence of Properties 5.1 and 5.2, one can easily show using Brouwer's Fixed Point Theorem (as in [Jia et al., 2015]), that there exists at least one fixed point $\mathbf{x}^* = F(\mathbf{x}^*)$ in the convex compact set \mathcal{A} . In [Jia et al., 2015], a method involving multiple complicated inequalities is used to show that the fixed point \mathbf{x}^* is unique. This is done separately to the convergence proof. In the following result, one is able to establish exponential convergence, and as a consequence of the method used, immediately prove that it is to a unique fixed point.*

In the following proof, a property is said to hold uniformly if the property holds for all $\mathbf{x}(s) \in \mathcal{A}$, where \mathcal{A} was defined in Property 5.2.

Theorem 5.2. *Suppose that $n \geq 3$ and suppose further that the network $\mathcal{G}[C]$ satisfies Assumption 5.1. Suppose further that $\mathcal{G}[C]$ does not have star topology. Then, the system Eq. (5.4), with initial conditions $\mathbf{x}(0) \in \Delta_n$, converges exponentially fast to a unique equilibrium point $\mathbf{x}^* \in \text{int}(\Delta_n)$.*

Proof. Consider any given initial condition $\mathbf{x}(0) \in \tilde{\Delta}_n$. According to Property 5.2, $\mathbf{x}(s) \in \mathcal{A}, \forall s > 0$ for a sufficiently small r . It remains to study the system Eq. (5.4) for $\mathbf{x}(s) \in \mathcal{A}$; in the following analysis, it is assumed that $s > 0$. The reader should review the concepts and terminology of Appendix A.3, which will be heavily utilised.

Define the Jacobian of F evaluated at \mathbf{x} as $J_{F(\mathbf{x})} = \{\frac{\partial F_i}{\partial x_j}(\mathbf{x})\}$. One obtains

$$\frac{\partial F_i}{\partial x_i}(\mathbf{x}) = \frac{\gamma_i \alpha(\mathbf{x})}{(1 - x_i)^2} - \frac{\gamma_i^2 \alpha(\mathbf{x})^2}{(1 - x_i)^3} \quad (5.6)$$

$$\frac{\partial F_i}{\partial x_j}(\mathbf{x}) = -\frac{\gamma_i \gamma_j \alpha(\mathbf{x})^2}{(1 - x_i)(1 - x_j)^2}, \quad j \neq i. \quad (5.7)$$

From Eq. (5.4) and Eq. (5.5), one is able to readily verify that the Jacobian entries, as given above, can be expressed as

$$\begin{aligned} \frac{\partial F_i}{\partial x_i}(\mathbf{x}(s)) &= x_i(s+1) \frac{1 - x_i(s+1)}{1 - x_i(s)} \\ \frac{\partial F_i}{\partial x_j}(\mathbf{x}(s)) &= -\frac{x_i(s+1)x_j(s+1)}{1 - x_j(s)}, \quad j \neq i, \end{aligned}$$

which establishes (via a slight abuse in notation) the relation between the Jacobian $J_{F(x(s))}$ evaluated at $x(s)$ and the social power vector $x(s+1)$. The reason for doing so will become clear shortly. Accordingly, this yields the following virtual dynamics

$$\delta x(s+1) = J_{F(x(s))} \delta x(s).$$

Note that $J_{F(x(s))}$ is well defined and uniformly continuous because $x_i(s) < 1 - r, \forall i, s$ for some sufficiently small $r > 0$. One can now proceed to use nonlinear contraction analysis. Specifically, consider the transformed virtual displacement

$$\delta z(s) = \Theta(x(s)) \delta x(s), \quad (5.8)$$

where $\Theta(x(s)) = \text{diag}((1 - x_i(s))^{-1})$, i.e. Θ is dependent on the state $x(s)$ but not dependent explicitly on s . Property 5.2 establishes that $1 > 1 - x_i(s) > r > 0$, which in turn implies that $\Theta(x(s))$ is uniformly nonsingular, with eigenvalues $\lambda_{\min}(\Theta(x(s))) > 1$ and $\lambda_{\max}(\Theta(x(s))) < 1/r$. In other words, $\kappa < \|\Theta(x(s))\|_1 < \kappa^{-1}$ for some $\kappa > 0, \forall x(s) \in \mathcal{A}$, as required (see Appendix A.3).

The transformed virtual dynamics is given by

$$\begin{aligned} \delta z(s+1) &= \Theta(x(s+1)) J_{F(x(s))} \Theta(x(s))^{-1} \delta z(s) \\ &= \bar{H}(x(s)) \delta z(s). \end{aligned} \quad (5.9)$$

where $\bar{H}(x(s)) = \Theta(F(x(s))) J_{F(x(s))} \Theta(x(s))^{-1}$ is the Jacobian associated with the transformed virtual dynamics. By denoting $\bar{\Phi}(x(s)) = J_{F(x(s))} \Theta(x(s))^{-1}$, one can write $\bar{H}(x(s)) = \Theta(F(x(s))) \bar{\Phi}(x(s))$. The matrix $\bar{\Phi}(x(s))$ is computed in Eq. (5.10) (see page 91), and note that it can be considered as being solely dependent on $x(s+1) = F(x(s))$. Therefore, define $\Phi(x(s+1)) = \bar{\Phi}(x(s))$. For brevity, the argument $x(s+1)$ is dropped where there is no ambiguity and one simply writes Φ .

For each row i , one has $\phi_{ii} = x_i(s+1)(1 - x_i(s+1))$ and $\phi_{ij} = -x_i(s+1)x_j(s+1)$ where ϕ_{ij} is the $(i, j)^{\text{th}}$ element of Φ . Because $s > 0$, Properties 5.2 and 5.3 establish that $0 < x_i(s) < 1 - r, \forall i$. It follows that all diagonal entries of Φ are uniformly strictly positive and all off-diagonal entries of Φ are uniformly strictly negative. Notice that $\Phi = \Phi^\top$. Lastly, for any row i , observe that there holds

$$\sum_{j=1}^n \phi_{ij} = x_i(s+1) \left[1 - x_i(s+1) - \sum_{j=1, j \neq i}^n x_j(s+1) \right] = 0$$

because $x_i(s+1) + \sum_{j=1, j \neq i}^n x_j(s+1) = 1$. In other words, Φ has row and column sums equal to 0. One can thus conclude that Φ is the weighted Laplacian⁷ associated with an undirected, *completely connected*⁸ graph with edge weights which vary with $x(s+1)$. The edge weights, $-\phi_{ij}$, are uniformly lower bounded away from zero and upper bounded away from 1. This implies that $0 = \lambda_1(\Phi) < \lambda_2(\Phi) \leq \dots \leq \lambda_n(\Phi) <$

⁷See Appendix A.2 for the definition and spectral properties of the Laplacian matrix of a graph.

⁸By completely connected, the author means that there is an edge going from every node i to every other node j .

$$\begin{aligned}
\bar{\Phi}(\mathbf{x}(s)) &= \begin{bmatrix} x_1(s+1) \frac{1-x_1(s+1)}{1-x_1(s)} & -\frac{x_1(s+1)x_2(s+1)}{1-x_2(s)} & \dots & -\frac{x_1(s+1)x_n(s+1)}{1-x_n(s)} \\ -\frac{x_1(s+1)x_2(s+1)}{1-x_1(s)} & x_2(s+1) \frac{1-x_2(s+1)}{1-x_2(s)} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{x_1(s+1)x_n(s+1)}{1-x_1(s)} & -\frac{x_2(s+1)x_n(s+1)}{1-x_2(s)} & \dots & x_n(s+1) \frac{1-x_n(s+1)}{1-x_n(s)} \end{bmatrix} \times \begin{bmatrix} 1-x_1(s) & & & \\ & 1-x_2(s) & & \\ & & \ddots & \\ & & & 1-x_n(s) \end{bmatrix} \\
&= \begin{bmatrix} x_1(s+1)(1-x_1(s+1)) & -x_1(s+1)x_2(s+1) & \dots & -x_1(s+1)x_n(s+1) \\ -x_1(s+1)x_2(s+1) & x_2(s+1)(1-x_2(s+1)) & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ -x_1(s+1)x_n(s+1) & -x_2(s+1)x_n(s+1) & \dots & x_n(s+1)(1-x_n(s+1)) \end{bmatrix}. \tag{5.10}
\end{aligned}$$

∞ , i.e., Φ is uniformly positive semidefinite with a single eigenvalue at 0, with the associated eigenvector $\mathbf{1}_n$.

Since $\bar{\Phi}(x(s)) = \Phi(x(s+1))$ and $\Theta(x(s+1)) = \Theta(F(x(s)))$, $\bar{H}(x(s))$ can be considered as depending solely on $x(s+1)$. Letting $H(x(s+1)) = \bar{H}(x(s))$, the calculation $H(x(s+1)) = \Theta(x(s+1))\Phi(x(s+1))$ yields that, for any $i \in \mathcal{I}$,

$$\begin{aligned} h_{ii}(x(s+1)) &= x_i(s+1) \\ h_{ij}(x(s+1)) &= -\frac{x_i(s+1)x_j(s+1)}{1-x_i(s+1)}, \quad j \neq i, \end{aligned}$$

where $h_{ij}(x(s+1))$ is the $(i, j)^{th}$ element of $H(x(s+1))$. For brevity, and when there is no risk of confusion, the argument $x(s+1)$ is dropped, writing H . Observe that the diagonal entries and off-diagonal entries of $H(x(s+1))$ are uniformly strictly positive and uniformly strictly negative, respectively. Notice that $\Phi\mathbf{1}_n = \mathbf{0}_n \Rightarrow H\mathbf{1}_n = \Theta(x(s+1))\Phi(x(s+1))\mathbf{1}_n = \mathbf{0}_n$. In other words, each row of H sums to zero. It follows that H is the weighted Laplacian matrix associated with a directed, *completely connected* graph with edge weights which vary with $x(s+1)$. The edge weights, $-h_{ij}$, are uniformly upper bounded away from infinity and lower bounded away from zero. It follows that the associated Laplacian matrix H has a single eigenvalue at 0, and all other eigenvalues have positive real parts (see Appendix A.2).

With $\Theta(x(s+1))$ uniformly symmetric and positive definite, and $\Phi(x(s+1))$ uniformly symmetric and positive semidefinite, it follows from Lemma A.4 that all eigenvalues of $H = \Theta\Phi$ are real and nonnegative. Combining with the above analysis, it follows that H has a single zero eigenvalue and *all other eigenvalues are strictly positive and real*. Observing that $\text{trace}(H) = \sum_{i=1}^n x_i(s+1) = 1 = \sum_{i=1}^n \lambda_i(H)$, yields $\max_i (\lambda_i(H)) < 1$ uniformly, since $n \geq 3$.

The stronger result, that $\|H\|_1 < 1$ uniformly, is now established. Because \mathcal{A} is compact, observe that $\|H\|_1 < 1$ if and only if, for all $i \in \mathcal{I}$, there holds $\sum_{j=1}^n |h_{ji}| < 1$. Equivalently,

$$x_i + \sum_{j=1, j \neq i}^n \left(\frac{x_i}{1-x_j} \right) x_j < 1, \quad (5.11)$$

and notice that the time argument $s+1$ has been dropped for brevity. From the fact that $x_i > 0, \forall i$ (recall Property 5.3 and that one is considering $s > 0$), and $n \geq 3$, it follows that $x_i + x_j < 1 \Rightarrow x_i/(1-x_j) < 1$ for all $j \neq i$. Combining this with the fact that $x_i + \sum_{j=1, j \neq i}^n x_j = 1$, one can immediately verify that Eq. (5.11) holds for all i . Because \mathcal{A} is bounded and compact, this implies that $\|H\|_1 < 1 - \eta$ for some $\eta > 0$ and all $x(s) \in \mathcal{A}$. Recalling the transformed virtual dynamics in Eq. (5.9) yields

$$\|\delta z(s+1)\|_1 = \|H(x(s+1))\delta z(s)\|_1 < (1-\eta)\|\delta z(s)\|_1.$$

It follows that the transformed virtual displacement δz converges to zero exponentially fast. Recall the definition of $\delta z(s)$ in Eq. (5.8), and the fact that $\Theta(x(s))$ is

uniformly nonsingular. It follows that $\delta x(s) \rightarrow \mathbf{0}_n$ exponentially, $\forall x(s) \in \mathcal{A}$.

The above results have established that \mathcal{A} is a *generalised contraction region* in accordance with Definition A.1. Because \mathcal{A} is compact and convex, it follows from Theorem A.4 and Corollary A.1 that *all trajectories* of $x(s+1) = F(x(s))$ with $x(0) \in \tilde{\Delta}_n$, converge exponentially fast to a *single trajectory*. According to Brouwer's Fixed Point Theorem [Khamsi and Kirk, 2011], there is at least one fixed point $x^* = F(x^*) \in \text{int}(\Delta_n)$, which is a trajectory of $x(s+1) = F(x(s))$. It then immediately follows that all trajectories of $x(s+1) = F(x(s))$ converge exponentially fast to a *unique fixed point* $x^* \in \text{int}(\Delta_n)$. \square

Having established a general convergence result, the following corollary presents some additional insight on the equilibria at the corners of the simplex Δ_n . As a by-product, a greatly simplified proof is obtained to show that F is continuous and smooth, which extends the result of Property 5.1.

Corollary 5.1 (Vertex Equilibrium). *For the system Eq. (5.4), the fixed point \mathbf{e}_i of the map $F(x)$ in Eq. (5.5) is unstable if $\gamma_i < 1/2$. If $\gamma_i = 1/2$, i.e., v_i is the centre node of a star graph, then the fixed point \mathbf{e}_i is asymptotically stable, but is not exponentially stable.*

Proof. Without loss of generality, consider $i = 1$. Observe that

$$F_1(x) = \frac{\gamma_1}{\gamma_1 + \sum_{k=2}^n \frac{(1-x_1)\gamma_k}{1-x_k}}$$

$$F_j(x) = \frac{\gamma_j(1-x_1)}{(1-x_j)(\gamma_1 + \sum_{k=2}^n \frac{\gamma_k(1-x_1)}{1-x_k})}, \forall j \neq 1,$$

and it is evident that these expressions are analytic in x_1 for all $x \in \Delta_n$. The same is then necessarily true of all their derivatives. It follows that F is continuous and smooth⁹ on Δ_n .

At $x = \mathbf{e}_1$, the expressions above yield that $F(\mathbf{e}_1) = \mathbf{e}_1$, and differentiating the expressions yields a value for the Jacobian at $x = \mathbf{e}_1$ in which $\frac{\partial F_1}{\partial x_1} = \frac{1-\gamma_1}{\gamma_1}$, $\frac{\partial F_i}{\partial x_1} = -\frac{\gamma_i}{\gamma_1}$, $\frac{\partial F_i}{\partial x_j} = 0$ for all $i, j \neq 1$. It follows that $J_{F(x)}$ has a single eigenvalue at $(1-\gamma_1)/\gamma_1$ and all other eigenvalues are 0. If $\gamma_1 < 1/2$, then $(1-\gamma_1)/\gamma_1 > 1$ and the fixed point \mathbf{e}_1 is unstable. If $\gamma_1 = 1/2$, then $J_{F(x)}$ has a single eigenvalue at 1. The discrete-time counterpart¹⁰ to [Khalil, 2002, Theorem 4.15] then rules out \mathbf{e}_1 as an *exponentially stable* fixed point of $F(x)$ (asymptotic stability was established in Lemma 5.1). \square

Remark 5.6. *During the analysis of H , it was first established that $\forall i$, the eigenvalues of $\lambda_i(H)$ are real, nonnegative and less than 1. This indicates that the trajectories of Eq. (5.4)*

⁹The proof in [Jia et al., 2015, Lemma 2.2] established continuity but not smoothness, and used a series of complex inequality calculations.

¹⁰[Khalil, 2002, Theorem 4.15] states that the equilibrium point $x = \mathbf{0}$ of the nonlinear continuous-time system $\dot{x} = f(t, x)$ is exponentially stable if and only if $x = \mathbf{0}$ is an exponentially stable equilibrium point of the linear system $\dot{x} = A(t)x$, where $A(t) = \frac{\partial f}{\partial x}(t, x)|_{x=\mathbf{0}}$. The proof of the discrete-time counterpart theorem follows the same steps as the proof of [Khalil, 2002, Theorem 4.15]

about \mathbf{x}^* are not oscillatory in nature. It also follows that $\rho(\mathbf{H}) < 1$. According to [Horn and Johnson, 2012, Lemma 5.6.10], there exists a submultiplicative matrix norm $\|\cdot\|'$ such that $\|\mathbf{H}\|' < 1$. However, it must be recalled that $\mathbf{H}(\mathbf{x}(s+1))$ is in fact a non-constant matrix which changes over the trajectory of the system Eq. (5.4). It is not immediately obvious, and in fact is not a consequence of the eigenvalue property, that a single submultiplicative matrix norm $\|\cdot\|''$ exists such that $\|\mathbf{H}\|'' < 1$ **for all** $\mathbf{x} \in \mathcal{A}$. Existence of such a norm $\|\cdot\|''$ would establish the desired stability property.

In fact, the system $\delta\mathbf{z}(s+1) = \mathbf{H}(\mathbf{x}(s+1))\delta\mathbf{z}(s)$, with $\mathbf{H} \in \mathcal{M}$, $\mathcal{M} = \{\mathbf{H}(\mathbf{x}(s+1)) : \mathbf{x}(s+1) \in \mathcal{A}\}$, can be considered as a discrete-time linear switching system with state $\delta\mathbf{z}$, and thus under arbitrary switching, the system is stable if and only if the **joint spectral radius** is less than 1, that is $\rho(\mathcal{M}) = \lim_{k \rightarrow \infty} \max_i \{\|\mathbf{H}_{i_1} \dots \mathbf{H}_{i_k}\|^{1/k} : \mathbf{H}_{i_j} \in \mathcal{M}\} < 1$ [Blondel and Tsitsiklis, 2000]. This is of course a more restrictive condition than simply requiring that $\rho(\mathbf{H}_i) < 1$. It is known that even when \mathcal{M} is finite, computing $\rho(\mathcal{M})$ is NP-hard [Tsitsiklis and Blondel, 1997] and the question “ $\rho(\mathcal{M}) \leq 1$?” is an undecidable problem [Blondel and Tsitsiklis, 2000]. The problem is made even more difficult because in this instant, the set \mathcal{M} is not finite. This motivated the search for the stronger, and nontrivial, result that $\|\mathbf{H}\|_1 < 1, \forall \mathbf{x} \in \mathcal{A}$.

Remark 5.7. For $\delta\mathbf{z}$ given in Eq. (5.8), one can integrate to obtain $z_i = -\ln(1 - x_i)$ where z_i is the i^{th} element of $\mathbf{z}(\mathbf{x})$. However, the above convergence arguments were not presented by first defining $\mathbf{z}(\mathbf{x})$ and then seeking to study $\mathbf{z}(s+1) = \mathbf{G}(\mathbf{z}(s))$. This is because the proof arose from considering $\mathbf{x}(s+1) = \mathbf{F}(\mathbf{x}(s))$ using nonlinear contraction analysis, which studies stability via differential concepts. It was through Eq. (5.8) that one integrated¹¹ to obtain $z_i = -\ln(1 - x_i)$. Moreover, it will be observed in Chapter 6 that by conducting analysis on the transformed Jacobian using nonlinear contraction theory, one can straightforwardly deal with issue-varying $\mathbf{C}(s)$.

Remark 5.8. Simulations showed that sometimes $|\lambda_{\max}(\mathbf{J}_F)| > 1$, which implies that it is not always possible to find a matrix norm $\|\cdot\|'$ such that $\|\mathbf{J}_{F(\mathbf{x}(s))}\|' < 1$ uniformly. Such a result would allow application of a standard method for proving Banach’s Contraction Theorem on the map \mathbf{F} [Khamisi and Kirk, 2011]. This is what motivated an approach using nonlinear contraction analysis, with transformation Θ . The work [Lohmiller and Slotine, 1998] specifically discusses contraction in the Euclidean metric $\|\delta\mathbf{z}\|_2 = \|\Theta\delta\mathbf{x}\|_2$. One requires $\lambda_{\max}(\mathbf{H}(\mathbf{x})^\top \mathbf{H}(\mathbf{x})) < 1$ to hold uniformly to guarantee that $\|\delta\mathbf{z}(s)\|_2 \rightarrow 0$ exponentially fast. However, simulations showed that $\lambda_{\max}(\mathbf{H}(\mathbf{x})^\top \mathbf{H}(\mathbf{x}))$ was often much larger than 1. This motivated the consideration of contraction in the absolute sum metric, with appropriate adjustments to the proof presented in Appendix A.3. Such an approach is alluded to in [Lohmiller and Slotine, 1998, Section 3.7].

Having established an exponential convergence result, study is now undertaken to obtain additional results on the dynamics of the system Eq. (5.4), which in turn reveal insights about the evolution of social power in the context of a social network.

¹¹In general, the entries of Θ may have expressions which do not have analytic antiderivatives, and thus an analytic $\mathbf{z}(\mathbf{x}(s), s)$ cannot always be found, but $\delta\mathbf{z}(s)$ can always be defined.

5.3 Further Analysis of Dynamical Behaviour

In this section, a result is first established that significantly expands Property 5.2 and introduces a stronger *contraction-like result* that is also applicable to social networks with star topology, unlike Property 5.2. This contraction-like property is used to establish two further results. The first is an upper bound on individual i 's social power at equilibrium x_i^* , as a function of γ_i , the i^{th} entry of the dominant left eigenvector γ^\top of C . The second result is on the convergence rate to x^* for a class of C .

5.3.1 A Contraction-Like Property

Recall Property 5.2, which established that if the system Eq. (5.4) was close to the boundary of the simplex Δ_n , then there was a contraction towards the centre of the simplex. The following theorem gives an explicit value for the region on which this contraction-like property holds, which for x_j depends only on γ_j .

Theorem 5.3. Suppose that $n \geq 3$, $x(0) \in \tilde{\Delta}_n$, and the social network $\mathcal{G}[C]$ satisfies Assumption 5.1. Define

$$r_j = \frac{1 - 2\gamma_j}{1 - \gamma_j}, \quad (5.12)$$

where γ_j is the j^{th} entry of γ^\top , the dominant left eigenvector of C . If $\gamma_j < 1/2$, which implies that $r_j > 0$, then for any $0 < r \leq r_j$, there holds

$$x_j = 1 - r \Rightarrow F_j(x) < 1 - r, \quad (5.13)$$

where $F_j(x)$ is the j^{th} entry of $F(x)$ given in Eq. (5.5). If $\gamma_j = 1/2$, then $\nexists r > 0 : r \leq r_j$, and thus Eq. (5.13) does not hold.

Proof. Property 5.3 established that for $x(0) \in \tilde{\Delta}_n$, there holds $x_i(s) > 0, \forall i$ and $s > 0$. Consider then $s > 0$. Since $x_j = 1 - r$ and $\gamma_j < 1/2$, then with $r \leq r_j$, there holds

$$\begin{aligned} F_j(x) &= \alpha(x) \frac{\gamma_j}{1 - x_j} = \frac{1}{\frac{\gamma_j}{1 - x_j} \left(1 + \frac{\sum_{k \neq j}^n \gamma_k / (1 - x_k)}{\gamma_j / (1 - x_j)}\right)} \frac{\gamma_j}{1 - x_j} \\ &= \frac{1}{1 + \frac{\sum_{k \neq j}^n \gamma_k / (1 - x_k)}{\gamma_j / (1 - x_j)}} = \frac{1}{1 + \sum_{k \neq j}^n \frac{r}{\gamma_j} \frac{\gamma_k}{(1 - x_k)}}. \end{aligned} \quad (5.14)$$

Because $1 - x_k < 1 \forall k \in \mathcal{I}$, one obtains $\gamma_k / (1 - x_k) > \gamma_k$, which implies that the right hand side of Eq. (5.14) obeys

$$\frac{1}{1 + \sum_{k \neq j}^n \frac{r}{\gamma_j} \frac{\gamma_k}{(1 - x_k)}} < \frac{1}{1 + \sum_{k \neq j}^n \frac{\gamma_k r}{\gamma_j}} \quad (5.15)$$

$$= \frac{1}{1 + \frac{(1 - \gamma_j)r}{\gamma_j}} = \frac{\gamma_j}{\gamma_j + (1 - \gamma_j)r}, \quad (5.16)$$

with the first equality obtained by noting that $\sum_{k \neq j}^n \gamma_k = 1 - \gamma_j$ according to the definition of γ . It follows from Eq. (5.14) and Eq. (5.16) that

$$\begin{aligned} 1 - r - F_j(\mathbf{x}) &> 1 - r - \frac{\gamma_j}{\gamma_j + (1 - \gamma_j)r} \\ &= \frac{\gamma_j + (1 - \gamma_j)r - r\gamma_j - (1 - \gamma_j)r^2 - \gamma_j}{\gamma_j + (1 - \gamma_j)r} = \frac{r(1 - 2\gamma_j) - r^2(1 - \gamma_j)}{\gamma_j + (1 - \gamma_j)r} \\ &= \frac{r(1 - \gamma_j) \left[\frac{1 - 2\gamma_j}{1 - \gamma_j} - r \right]}{\gamma_j + (1 - \gamma_j)r}. \end{aligned}$$

Substituting in r_j from Eq. (5.12) then yields

$$1 - r - F_j(\mathbf{x}) > \frac{r(1 - \gamma_j)(r_j - r)}{\gamma_j + (1 - \gamma_j)r} \geq 0 \quad (5.17)$$

because $r_j \geq r$. In other words, $1 - r > F_j(\mathbf{x})$, which completes the proof. \square

This contraction-like result is now used to establish an upper bound on the social power of an individual at equilibrium.

5.3.2 Upper Bound on Individual's Social Power at Equilibrium

Theorem 5.1, Property (iii) established that the social power at equilibrium \mathbf{x}^* is dependent only on the network topology, C , and specifically the dominant left eigenvector γ^\top of C . Property (ii) of Theorem 5.1 provides an ordering result of x_i^* and x_j^* . However, no general results appear to exist for analytical computation of the vector \mathbf{x}^* given γ^\top . Results exist for some special cases, though, such as for doubly stochastic C and for \mathcal{G} with star topology [Jia et al., 2015]. While no explicit equality relating x_i^* to γ_i is provided, an explicit *inequality* is provided.

Corollary 5.2 (Upper bound on x_i^*). *Suppose that $n \geq 3$ and $\mathbf{x}(0) \in \tilde{\Delta}_n$. Suppose further that $\mathcal{G}[C]$ satisfies Assumption 5.1, and is not a star graph. Then, the i^{th} entry of the unique fixed point \mathbf{x}^* of F , given in Eq. (5.5), satisfies $x_i^* < \gamma_i / (1 - \gamma_i)$.*

Proof. Theorem 5.3 established that, for any $j \in \mathcal{I}$, if $x_j = 1 - r$ with $r \leq r_j$, then the system Eq. (5.4) will contract so that $F_j(\mathbf{x}(s)) < x_j$. It is straightforward to conclude that the map $F(\mathbf{x})$ continues to contract towards the centre of the simplex Δ_n until $x_i(s) < 1 - r_i, \forall i$, where r_i is given by Eq. (5.12).

Suppose that $x_j^* \geq 1 - r_j = \gamma_j / (1 - \gamma_j)$. According to the arguments in the paragraph above, there holds $F_j(\mathbf{x}^*) < 1 - r_j \leq x_j^*$. On the other hand, the definition of \mathbf{x}^* as a fixed point of F implies that $x_j^* = F_j(\mathbf{x}^*)$, which leads to a contradiction. Therefore, $x_j^* < 1 - r_j = \gamma_j / (1 - \gamma_j)$ as claimed. \square

Note that this result is separate from the result of Theorem 5.2, which deals with convergence to \mathbf{x}^* . Here, an upper bound is established for the values of the entries of

the unique fixed point \mathbf{x}^* , i.e., the social power at equilibrium, given γ . In particular, the bound for individual i 's social power at equilibrium depends only on γ_i .

Remark 5.9. *Three specific conclusions can be drawn from these two results. First, in relation to the transient behaviour of $\mathbf{x}(s)$, Theorem 5.3 indicates that for any i and with $x_i(0)$ small, the peak overshoot of $x_i(s)$ above x_i^* is bounded as $x_i(s) \leq \gamma_i/(1 - \gamma_i)$. Second, suppose that $\mathcal{G}[\mathbf{C}]$ has star topology with centre node v_1 . Then, $\gamma_1 = 0.5$ according to Fact 5.1, and thus x_1 does not decrease according to Theorem 5.3. This is consistent with the findings in [Jia et al., 2015], i.e., Lemma 5.1. Last, suppose that \mathcal{G} is strongly connected and that $\gamma_i < 1/3, \forall i \in \mathcal{I}$. Then, according to Corollary 5.2, no individual will have more than half of the total social power at equilibrium, i.e., $x_i^* < 1/2, \forall i \in \mathcal{I}$. This provides a sufficient condition on the social network topology to ensure that no individual has a dominating presence in the opinion discussion.*

Remark 5.10. *[Tightness of the Bound] The tightness of the bound $x_i^* < \gamma_i/(1 - \gamma_i)$ increases as γ_k decreases $\forall k \neq i$. This is in the sense that the ratio $x_i^*(1 - \gamma_i)/\gamma_i$ approaches 1 from below as γ_k decreases $\forall k \neq i$. This conclusion can be drawn by noting that in order to obtain Eq. (5.15), the inequality $1 - x_k < 1$ is employed. From the fact that $\lim_{x_k \rightarrow 0} 1 - x_k = 1$, and because the contraction-like property of Theorem 5.3 holds for $x_k \geq \gamma_k/(1 - \gamma_k)$, it follows that the tightness of the bound $x_i^* < \gamma_i/(1 - \gamma_i)$ increases as γ_k decreases $\forall k \neq i$. If there is a single individual i with $\gamma_i \gg \gamma_k, \forall k \neq i$, one can accurately estimate x_i^* , e.g. if $\gamma_i > 1/3 \gg \gamma_k, \forall k \neq i$, it can be said with high confidence that individual i will hold more than half of the total social power at equilibrium, i.e., $x_i^* > 0.5$.*

5.3.3 Convergence Rates

A result is now presented bounding the convergence rate for a \mathbf{C} belonging to an important subset of all possible \mathbf{C} matrices; in this subset, no individual i will have more than half the social power at equilibrium, i.e. $\nexists i : x_i^* \geq 0.5$.

Lemma 5.2 (Convergence Rate). *Suppose that $\mathbf{C} \in \mathcal{L}$, where $\mathcal{L} = \{\mathbf{C} \in \mathbb{R}^{n \times n} : \gamma_i < 1/3, \forall i, n \geq 3\}$ ¹² and γ_i is the i^{th} entry of the dominant left eigenvector γ^\top associated with \mathbf{C} . Then, with $\mathbf{x}(0) \in \tilde{\Delta}_n$, there exists a finite s_1 such that for all $s \geq s_1$, the system Eq. (5.4) contracts to its unique equilibrium point \mathbf{x}^* with a convergence rate obeying*

$$\|\mathbf{x}^* - \mathbf{x}(s+1)\|_1 \leq (2\beta - \epsilon)\|\mathbf{x}^* - \mathbf{x}(s)\|_1 \quad (5.18)$$

where $2\beta - \epsilon < 1$, with $\beta = \max_{i \in \mathcal{I}} \gamma_i/(1 - \gamma_i) < 1/2$ and ϵ being a sufficiently small positive constant.

Proof. From Theorem 5.3, one concludes that $x_i(s) \geq 1 - \beta \Rightarrow F_i(\mathbf{x}(s)) = x_i(s+1) < x_i(s)$, where $\beta = \max_i \beta_i$. This means that there exists a sufficiently small positive constant ϵ and $s_1 < \infty$ such that $x_i(s) \leq \beta - \epsilon$ for all $s \geq s_1$. (That $s_1 < \infty$ is a consequence of the compactness of $\tilde{\Delta}_n$). In other words, $x_i(s)$ will be no greater than

¹²According to Fact 5.1, \mathcal{L} does not contain any \mathbf{C} whose associated graph has a star topology.

$\beta - \epsilon/2$ in a finite number of issues after $s = 0$; this fact will be used below to upper bound the norm of the untransformed Jacobian.

Each column sum of the Jacobian $J_{F(x(s))}$ is equal to zero, which comes from the fact that $\sum_i F_i(x) = 1$ (one can verify this property using the expressions in Eq. (5.6) and Eq. (5.7)). Note also that the diagonal entries of the Jacobian are strictly positive and for $s \geq s_1$, there holds $\partial F_i / \partial x_i \leq \beta - \epsilon/2$, $\forall i$. This is because $x_i(1 - x_i) \leq (\beta - \epsilon/2)(1 - \beta + \epsilon/2)$ for $x_i \leq \beta - \epsilon/2 < 0.5$ and $1/(1 - x_i) \leq 1/(1 - \beta + \epsilon/2)$. Combining the column sum property and the fact that the off-diagonal entries of the Jacobian are strictly negative, it follows that for $s \geq s_1$, there holds $\|J_{F(x(s))}\|_1 = 2 \max_i \partial F_i / \partial x_i \leq 2\beta - \epsilon < 1 - \eta$ where η is an arbitrarily small positive constant. The inequality in Eq. (5.18) follows immediately. \square

Remark 5.11. The quantity $2\beta - \epsilon$ (which is a Lipschitz constant associated with the map F) is an upper bound on the convergence rate of the system as in Eq. (5.18). By assuming $\gamma_i < 1/3$, $\forall i$, one is able to work directly with the Jacobian J_F , as opposed to the transformed Jacobian H . It is in general much more difficult to compute an upper bound on $\|H\|_1$ using γ and Corollary 5.2 when $\exists i : \gamma_i \geq 1/3$. Note that \mathcal{L} includes many of the topologies likely to be encountered in social networks. Topologies for which $\exists i : \gamma_i \geq 1/3$ will have an individual who holds more than half the social power at equilibrium. Such topologies are more reflective of autocracy-like or dictatorship-like networks, as opposed to a group of nominally equal peers.

5.4 Conclusions

This chapter revisited the recently proposed DeGroot–Friedkin model, with constant network topology. The model considers how an individual’s social power evolves as the influence network discusses a sequence of issues. A number of extensions have been presented on the original model. The key contribution is the employment of nonlinear contraction analysis to establish an exponential convergence result for general topologies, whereas the original paper [Jia et al., 2015] only established an asymptotic convergence result. Unstable equilibria have also characterised. For special star-topology networks, it has been proved that the stable equilibrium was asymptotically stable, but was not exponentially so. The second major development was to extend an existing “contraction-like” property, from which two further results were established; an upper bound on an individual’s social power at equilibrium, and the convergence rate of the model for a general class of network topologies.

Dynamic Social Networks: Exponential Forgetting of Perceived Social Power

Chapter Summary

This chapter extends the DeGroot–Friedkin model by introducing dynamic network topology, and in particular, issue-varying relative interaction topology (the term “issue” is used interchangeably with “topic”). The dynamic network topology problem is formally defined, with motivating examples provided as to why issue-varying networks are reflective of real-world social networks. Then, the nonlinear contraction analysis framework introduced in Chapter 5 is used to draw a key conclusion: under mild assumptions on the properties of the issue-varying network, each individual’s initial (perceived) social power is forgotten exponentially fast, and in the limit of the issue sequence, each individual’s social power converges to a “unique limiting trajectory” that depends only on the issue-varying network topology. The results in Section 5.3 concerning the upper bound on an individual’s social power at equilibrium and convergence rate for a class of topologies are also extended, with obvious modifications due to the dynamic topology. As a special case, periodically-varying topologies are also considered.

6.1 Introduction

In real-world social networks, the network topology is almost always evolving over time. New interpersonal relationships may be formed, old ones might be removed, or the strength of the interaction may change. This gives is a strong incentive to study opinion dynamics models where the topology is time-varying or dynamic.

Roughly speaking, dynamic topology can fall into at least two broad categories. The first is where the topology varies explicitly as a function of time, e.g. the influence weight $w_{ij}(t)$ in the DeGroot or Friedkin–Johnsen models. In such studies, a typical objective is to obtain *graphical conditions* on the topology which guarantee a certain phenomenon is achieved. For example, a “repeatedly jointly strongly

connected” condition on the switching of $\mathcal{G}[\mathbf{W}(t)]$ is a sufficient condition for the time-varying DeGroot model

$$\mathbf{y}(t+1) = \mathbf{W}(t)\mathbf{y}(t) \quad (6.1)$$

to reach a consensus of opinions exponentially fast. A number of results exist, with some works stating equivalent graphical conditions using different terminology or definitions. Thus, no attempt will be made to provide a detailed comparison and study of these results. For the DeGroot model, time-varying $\mathcal{G}[\mathbf{W}(t)]$ has been studied in e.g. [Ren and Beard, 2005; Cao et al., 2008; Jadbabaie et al., 2003; Nedić and Liu, 2017; Shi and Johansson, 2013]. The continuous-time Abelson model has also been heavily studied on time-varying networks, because of its relevance to distributed control of multi-agent systems (see Section 1.2.5). Example works include [Ren and Beard, 2005; Anderson et al., 2017; Shi and Johansson, 2013; Mesbahi and Egerstedt, 2010a; Ren and Beard, 2007]. The time-varying Altafini model has been studied in discrete-time [Liu et al., 2017a] and in continuous-time [Proskurnikov et al., 2016]. These works consider topology that is explicitly dependent on time, and may be interested in establishing the conditions for which the *same phenomenon* arises in time-varying networks as in time-invariant networks.

The second category is where the topology changes in a *state-dependent* manner, and since the state (typically the individuals’ opinions in an opinion dynamics model) often evolves over time (before perhaps reaching a steady state in some models), then necessarily the network topology is dynamic, though this thesis will refrain from calling these *time-varying* networks since the network is not an explicit function of time. For example, the influence weight $w_{ij}(\mathbf{y}(t))$ of the DeGroot model may depend on the values of opinion $y_i(t), y_j(t)$. A large number of works fall into this category, including the Hegselmann–Krause model (and its many variants), and the works of e.g. [Mäs et al., 2014; Duggins, 2017; Dandekar et al., 2013; Amelkin et al., 2017]. A key aspect of such works is to establish proper motivation of the process by which the network topology evolves. Almost certainly, the functional dependence of an influence weight on the opinions, given as $w_{ij} = f_i(\mathbf{y})$, must be carefully modelled to reflect established sociological and social psychological concepts such as homophily, bias assimilation, pressure to conform, etc. It is often the case that the initial opinion values and the functional dependence (including parameters in the function) determines the limiting behaviour (assuming the opinions exist in the limit), and there is no guarantee that the network remains connected. For example, the number of clusters, and distance between clusters in the Hegselmann–Krause model is linked to the size of the confidence bound on each individual. Moreover, each cluster is disconnected from the others. Attempts to study such networks may be undertaken only with help of thorough examination of supporting literature from the social sciences (including phenomena arising from the social processes).

It should now be clear why dynamic networks are of interest in agent-based opinion dynamics models. With respect to the DeGroot–Friedkin model, the influence weights already vary between the issues $s \in \mathcal{S}$, i.e. $w_{ij}(s)$, via the social process of

reflected self-appraisal. In this chapter, dynamic topology is therefore in reference to relative interaction matrices $C(s)$ which are different between issues, *but remain constant during the period of discussion for any given issue*. Relative interaction matrices encode trust or relationship strength between individuals in a network, and in Section 6.1.1 below, reasons are given as to why the topology might be dynamic, with examples provided in support. It is assumed that the relative interaction matrices do not vary in a manner dependent on the social powers of the individuals, but can otherwise vary arbitrarily. In other words, this chapter deals with the first category of dynamic networks, but as will be clarified in the sequel, the set of such networks within the DeGroot–Friedkin framework is large.

In this chapter, a general convergence result is explained in the context of the DeGroot–Friedkin model. In particular, it is discovered that the process of DeGroot opinion updating and reflected self-appraisal leads to a self-regulating network. That is, one can consider an individual's initial social power as the individual's own estimate of his/her social power before any discussion occurs (it can also be considered as his/her perceived power), and analysis shows that each individual *forgets his/her initial (i.e. perceived) social power exponentially fast*. The true social power of any individual at any given issue is determined only by the previously occurring sequence of dynamic relative interpersonal interaction topology.

6.1.1 Motivating Examples for Issue-Varying Topology

In the introduction above, qualitative reasons were provided as to why social networks might have dynamically changing topology, for a variety of models. Here, quantitative details on the problem of interest are provided. Having introduced the DeGroot–Friedkin model in the previous chapter, it is appropriate to expand on the motivation here with specific reference to the model (see Section 5.1.1). Specifically, examples are provided to show how relative interaction matrices $C(s)$ may vary dynamically between issues as a result of *issue-driven* or *individual-driven* topology change. Henceforth, reference to “dynamic” topology or $C(s)$ is a reference to $C(s)$ which varying between issues $s_i, s_j \in \mathcal{S}$, but remain constant for a given s .

Issue-driven: Consider a government cabinet that meets to discuss the issues of defence, economic growth, social security programs and foreign policy. Each minister (individual in the cabinet) has a specialist portfolio (e.g. defence) and perhaps a secondary portfolio (e.g. foreign policy). While every minister will partake in the discussion of each issue, the weights $c_{ij}(s)$ will change. For example, if minister i 's portfolio is on defence, then $c_{ji}(s_{\text{defence}})$ will be high as other ministers j place more trust on minister i 's opinion. On the other hand, $c_{ji}(s_{\text{social}})$ will be low. It is then apparent that $C(s_{\text{defence}}) \neq C(s_{\text{social}})$ in general. This motivates the incorporation of *issue-dependent* or *issue-driven* topology change into the DeGroot–Friedkin model.

Individual-driven: Dynamic relative interaction matrices are a natural way of describing *network structural changes over time*. For many reasons, relationships may form, change, or die out. As an example, consider individual i and individual j in a network, with $c_{ij}(0) > 0$, and suppose that $y_i(0, s) \gg y_j(0, s)$ for $s = 0, 1, \dots, 5$, i.e.

individual j consistently holds an initial opinion vastly different from individual i . Then, i may decide that j is not worth listening to, and set $c_{ij}(6) = 0$. This is similar to the concept of homophily as it appears in the Hegselmann–Krause model (see Section 1.2.2), where individuals interact only with others who hold similar views.

The two examples above are different from each other, but both equally provide motivation for *dynamic* topology. It is therefore of interest to consider the DeGroot–Friedkin model with dynamic $C(s)$ and to determine how such dynamic topology may affect or change an individual's social power evolution over the issue sequence.

6.1.2 Chapter Organization

In Section 6.2, the DeGroot–Friedkin model with dynamic topology is presented. Section 6.3 presents the main convergence result and extends a “contraction-like” property to include dynamic topology. Section 6.4 studies the special case of periodically-varying topology, and simulations are given in Section 6.5 to illustrate the main results. Conclusions are drawn in Section 6.6.

6.2 The Dynamic Topology Model and Objective

To begin, the DeGroot–Friedkin model is extended to consider dynamic topology, and an assumption is then placed on the properties of the dynamic $C(s)$.

First, the opinion dynamics for each issue is given by

$$y(s, t + 1) = W(s)y(s, t), \quad (6.2)$$

where

$$W(s) = X(s) + (I_n - X(s))C(s). \quad (6.3)$$

This records the fact that $C(s)$ varies between issues but is constant for all t for a given issue, in distinction to Eq. (5.2) in the previous chapter, where C was constant for each issue, and constant over the issue sequence. Precise details of the adjustments to the model arising from dynamic C are stated below. Here, an assumption is placed on the way in which $C(s)$ varies $\forall s \in \mathcal{S}$.

Assumption 6.1. *For a given sequence of $C(s)$, $s = 0, 1, 2, \dots$, the sequence is such that the entries of $C(s_1)$, given as $c_{ij}(s_1)$, do not depend on $x(s')$, $\forall s' \leq s_1$.*

Assumption 6.1 ensures that the dominant left eigenvector $\gamma^\top(s)$ is independent of the state $x(s)$, because $C(s)$ is independent of $x(s)$. Notice that both the issue-driven and individual-driven examples in Section 6.1.1 satisfy Assumption 6.1. Almost all issue-driven dynamics $C(s)$ satisfy the assumption because the sequence of $C(s)$ depends only on the sequence of issues. That is, for *analysis purposes* the sequence of $C(s)$ is considered to be determined before discussion begins on $s = 0$, but individual i may not necessarily know the sequence of $c_{ij}(s)$ a priori.

However, a situation where individual i adjusts $c_{ij}(s)$ to be larger because i observed that individual j had large impact $\zeta_j(s - 1)$, does not satisfy the assumption. For

social network models with state-dependent parameters, limiting behaviour depends critically on the function relating the parameters to the state (as detailed earlier). For social systems (as opposed to e.g. autonomous robotic systems), this functional dependence must necessarily reflect socio-psychological concepts. It is beyond the scope of this chapter to propose such functional dependence. Thus, Assumption 6.1 is in place, and investigation of how individuals might determine $C(s)$ based on $x(s), \dots, x(0)$ may give rise to interesting future work. Furthermore, Assumption 5.1 is also assumed to hold throughout this chapter, and it is restated here for the reader's convenience, with a minor adjustment to reflect the dynamic topology.

Assumption 6.2. *For all $s \in \mathcal{S}$, the graph $\mathcal{G}[C(s)] = (\mathcal{V}, \mathcal{E}[C(s)], C(s))$, with the same node set \mathcal{V} of $n \geq 3$ nodes, is strongly connected and no node v_i has a self-loop. Furthermore, the relative interaction matrix $C(s)$ row-stochastic for all $s \in \mathcal{S}$.*

The problem treated in this chapter is embedded in the following key objective.

Objective 6.1. *Consider a network $\mathcal{G}[C(s)]$ of $n \geq 3$ individuals that discuss a sequence of issues indexed by $s \in \mathcal{S} = \{0, 1, 2, \dots\}$. Suppose that Assumptions 6.1 and 6.2 hold. For each issue s , the DeGroot process Eq. (6.2) is used to describe the opinion dynamics, with the influence matrix $W(s)$ determined according to Eq. (6.3). Suppose that at the end of each issue, the vector of social power $x(s)$ defined above Eq. (5.2) changes according to Eq. (5.3). Then, the objective is to study the dynamical evolution (including convergence) of $x(s)$ over the sequence of issues \mathcal{S} , and determine what (if any) parameters of $\mathcal{G}[C(s)]$ affect the evolution of $x(s)$.*

For a population of $n \geq 3$ individuals, consider a social network $\mathcal{G}[C(s)] = (\mathcal{V}, \mathcal{E}[C(s)], C(s))$, with a finite set \mathcal{C} of P possible relative interaction matrices defined as $\mathcal{C} = \{C_p \in \mathbb{R}^{n \times n} : p \in \mathcal{P}\}$ where $\mathcal{P} = \{1, 2, \dots, P\}$. For simplicity, it is also assumed that $\nexists p$ such that the graph $\mathcal{G}[C_p]$ has star topology (see Definition 5.1). Let $\sigma(s) : \mathbb{N} \rightarrow \mathcal{P}$ be a switching signal, and suppose that $\sigma(s)$ determines the dynamic switching as $C(s) = C_{\sigma(s)}$. Suppose that Assumption 6.2 holds¹ for all $s \in \mathcal{S}$. Suppose further that Assumption 6.1 holds, which implies for any $s_1 \geq 0$, $\sigma(s_1)$ is independent of the state $x(s)$, for all $s < s_1$. Then, the DeGroot–Friedkin model with dynamic relative interaction matrices is

$$x(s+1) = F_{\sigma(s)}(x(s)), \quad (6.4)$$

where the nonlinear map $F_p(x)$ for $p \in \mathcal{P}$, is defined as

$$F_p(x) = \begin{cases} e_i & \text{if } x = e_i \text{ for any } i \\ \alpha_p(x) \begin{bmatrix} \frac{\gamma_{p,1}}{1-x_1} \\ \vdots \\ \frac{\gamma_{p,n}}{1-x_n} \end{bmatrix} & \text{otherwise} \end{cases}, \quad (6.5)$$

¹Assumption 6.2 is equivalent to requiring that Assumption 5.1 holds separately for all C_p , $p \in \mathcal{P}$.

where $\alpha_p(\mathbf{x}) = 1 / \sum_{i=1}^n \frac{\gamma_{p,i}}{1-x_i}$ and $\gamma_{p,i}$ is the i^{th} entry of the dominant left eigenvector of \mathbf{C}_p , $\gamma_p^\top = [\gamma_{p,1}, \gamma_{p,2}, \dots, \gamma_{p,n}]$. The derivation for Eq. (6.5) is an extension of the derivation of Eq. (5.5) using [Jia et al., 2015, Lemma 2.2], by noting that $\mathbf{C}(s) = \mathbf{C}_{\sigma(s)}$.

Remark 6.1. Analysis using the usual techniques for switched systems is difficult for the system Eq. (6.4). For arbitrary switching, one might typically seek to find a common Lyapunov function, i.e., one which would establish convergence for any fixed value of $p \in \mathcal{P}$. This, however, appears to be difficult (if not impossible) for Eq. (6.4). In the constant \mathbf{C} case studied in [Jia et al., 2015], the convergence result relied on 1) a Lyapunov function which was dependent on the unique equilibrium point \mathbf{x}^* , and 2) LaSalle's Invariance Principle. Both 1) and 2) are invalid when analysing Eq. (6.4). In the case of 1), the system Eq. (6.4) does not have a unique equilibrium point \mathbf{x}^* but rather a unique trajectory $\mathbf{x}^*(s)$ (as will be made clear in the sequel). In the case of 2), LaSalle's Invariance Principle is not applicable to general non-autonomous systems.

The following three properties were detailed in Chapter 5, and are detailed here again (with adjusted numbering) for the convenience of the reader.

Property 6.1. The map $F_p(\mathbf{x})$ in Eq. (6.5) is continuous on Δ_n for all $p \in \mathcal{P}$.

Property 6.2. For the system $\mathbf{x}(s+1) = F_p(\mathbf{x}(s))$ with map F_p given in Eq. (6.5), and $\mathcal{G}[\mathbf{C}_p]$ does not have star topology, there exists a sufficiently small r such that for any $r' \leq r$, $x_i(s) = 1 - r'$ implies $x_i(s+1) < 1 - r'$, for all i . Thus, for the set² $\mathcal{A} = \{\mathbf{x} \in \Delta_n : 1 - r \geq x_i \geq 0, \forall i \in \{1, \dots, n\}\}$, there holds $F_p(\mathcal{A}) \subset \mathcal{A}$.

Property 6.3. For the system $\mathbf{x}(s+1) = F_p(\mathbf{x}(s))$ with map F_p given in Eq. (6.5), if $\mathbf{x}(s_1) \in \tilde{\Delta}_n$ for some $s_1 < \infty$, then $\mathbf{x}(s) \in \text{int}(\Delta_n)$ for all $s > s_1$.

The DeGroot–Friedkin model with dynamic relative interpersonal interactions has now been formally defined, and Objective 6.1 is investigated in the remaining sections of this chapter.

6.3 Exponential Convergence to a Unique Limiting Trajectory

The main convergence result of this chapter is now stated, the proof of which turns out to be fairly straightforward. This is a consequence of the analysis framework arising from the techniques used in the proof of Theorem 5.2. In addition, it should be noted that in the theorem statement immediately below, a relaxation of the initial conditions is made; it is no longer required that $\sum_{i=1}^n x_i(0) = 1$. A social interpretation of the restrictive assumption $\sum_{i=1}^n x_i(0) = 1$ and the new, milder assumption, are given in Remark 6.2 immediately following the theorem, and an interpretation of the theorem itself is given in Remark 6.3.

²The sets Δ_n , $\tilde{\Delta}_n$, and $\text{int}(\Delta_n)$ were defined in Section 2.1, and are now redefined for the reader's convenience: The n -simplex is $\Delta_n = \{\mathbf{x} \in \mathbb{R}^n : 0 \leq \mathbf{x}, \mathbf{1}_n^\top \mathbf{x} = 1\}$, and $\tilde{\Delta}_n = \Delta_n \setminus \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ and $\text{int}(\Delta_n) = \{\mathbf{x} \in \mathbb{R}^n : 0 < \mathbf{x}, \mathbf{1}_n^\top \mathbf{x} = 1\}$.

Theorem 6.1. Suppose Assumptions 6.1 and 6.2 hold, and that $\nexists p$ such that $C_p \in \mathcal{C}$ is associated with a star topology graph, with \mathcal{C} defined above Eq. (6.4). Then, for system Eq. (6.4), a) there exists a unique trajectory $\mathbf{x}^*(s) \in \text{int}(\Delta_n), s \geq 0$ determined solely by $\gamma_{\sigma(s_1), s_1} \leq s$, and b) for all initial conditions satisfying $0 \leq x_i(0) < 1, \forall i \in \mathcal{I}$ and $\exists j \in \mathcal{I} : x_j(0) > 0$, there holds $\lim_{s \rightarrow \infty} [\mathbf{x}(s) - \mathbf{x}^*(s)] = \mathbf{0}_n$ exponentially fast, and c) if $\mathbf{x}(0) = \mathbf{e}_i$ for some i , then $\mathbf{x}(s) = \mathbf{e}_i, \forall s$.

Proof. First, regarding c), observe that if $\mathbf{x}(0) = \mathbf{e}_i$ for some i , then Eq. (6.5) leads to the conclusion that $\mathbf{x}(s) = \mathbf{e}_i$ for all s .

With initial conditions $x_i(0) < 1$, the map $F_{\sigma(0)}(\mathbf{x}(s)) \neq \mathbf{e}_i$ for any i . It is easy to verify that with these initial conditions, the matrix $W(0)$ is row-stochastic, irreducible and aperiodic, which implies that for $s = 0$, the opinions $\mathbf{y}(0, t)$ converge to a consensus as in the constant C case. Because $C(0)$ is irreducible, this implies that $\gamma_{\sigma(0), i} > 0$ for all i , and it follows that $\alpha_{\sigma(0)}(\mathbf{x}(0)) > 0$ because $\exists j : x_j(0) > 0$. One thus concludes that $\mathbf{x}(1) = F_{\sigma(0)}(\mathbf{x}(0)) > 0$, i.e., for issue $s = 1$, every individual's social power/self-weight is strictly positive, and the sum of the weights is 1. Moreover, because C_p is irreducible $\forall p$, this implies that for any p , there holds $\gamma_{p, i} > 0$ for all i . It follows that for $s \geq 1$, $\alpha_{\sigma(s)}(\mathbf{x}(s)) > 0$, which in turn guarantees that $\mathbf{x}(s+1) = F_{\sigma(s)}(\mathbf{x}(s)) > 0$, i.e., $\mathbf{x}(s) \in \text{int}(\Delta_n)$ for all $s > 0$. This ensures that $\forall s$, $W(s)$ is row-stochastic, and primitive. This implies that $\mathbf{y}(s, t)$ converges to a consensus for every issue $s \in \mathcal{S}$ (see Section 2.3.1) and thus the dominant left eigenvector $\zeta(s)^\top$ of $W(s)$, used to determine $\mathbf{x}(s+1) = \zeta(s)$, is always well defined.

Denote the i^{th} entry of F_p by $F_{p, i}$. Regarding Property 6.2, for each map F_p , define the set $\mathcal{A}_p(r_p) = \{\mathbf{x} \in \Delta_n : 1 - r_p \geq x_i \geq 0, \forall i \in \mathcal{I}\}$, where $0 < r_p \ll 1$ is sufficiently small such that $x_i(s) \leq 1 - r_p$ for all $i \in \mathcal{I}$ implies that $F_{p, i}(\mathbf{x}(s)) = x_i(s+1) < 1 - r_p$. Define $\bar{\mathcal{A}} = \{\mathbf{x} \in \Delta_n : 1 - \bar{r} \geq x_i \geq 0, \forall i \in \mathcal{I}\}$ where $\bar{r} = \min_p r_p$. Because $F_p(\bar{\mathcal{A}}) \subset \bar{\mathcal{A}}$, it follows that $\cup_{p=1}^P \mathcal{A}_p \subset \bar{\mathcal{A}}$, and that for the system Eq. (6.4), for all $s > 0$, $\mathbf{x}(s) \in \bar{\mathcal{A}}$, with $\bar{\mathcal{A}}$ compact and convex.

Denoting the Jacobian for the system Eq. (6.4) at issue s as $J_{F_{\sigma(s)}} = \{\frac{\partial F_{\sigma(s), i}}{\partial x_j}\}$, and from Assumption 6.1 yields that $\gamma_{\sigma(s)}$ is independent of $\mathbf{x}(s)$, one obtains

$$\frac{\partial F_{\sigma(s), i}}{\partial x_i}(\mathbf{x}) = \frac{\gamma_{\sigma(s), i} \alpha_{\sigma(s)}(\mathbf{x})}{(1 - x_i)^2} - \frac{(\gamma_{\sigma(s), i} \alpha_{\sigma(s)}(\mathbf{x}))^2}{(1 - x_i)^3}.$$

Similarly, for $j \neq i$, one obtains

$$\frac{\partial F_{\sigma(s), i}}{\partial x_j}(\mathbf{x}) = -\frac{\gamma_{\sigma(s), i} \gamma_{\sigma(s), j} (\alpha_{\sigma(s)}(\mathbf{x}))^2}{(1 - x_i)(1 - x_j)^2}.$$

Comparing to Eq. (5.6) and Eq. (5.7), one concludes that the Jacobian of the non-autonomous system Eq. (6.4) with the map given in Eq. (6.5) is expressible in the same form as the Jacobian of the original system Eq. (5.4) with the map given in Eq. (5.5). More precisely, it can be expressed in a form which is dependent on the trajectory of the system Eq. (6.4). Using the same transformation of $\delta \mathbf{z}$ given in

Eq. (5.8) with the same $\Theta(x(s))$, one obtains the exact same transformed virtual dynamics Eq. (5.9), expressed as

$$\delta z(s+1) = \mathbf{H}(x(s+1))\delta z(s). \quad (6.6)$$

Similar to the proof of Theorem 5.2, for some arbitrarily small $\eta > 0$, one can show that there holds $\|\mathbf{H}\|_1 < 1 - \eta$ for all $x(s) \in \bar{\mathcal{A}}$, independent of $p \in \mathcal{P}$, which implies that $\bar{\mathcal{A}}$ is a generalised contraction region. It follows that $\delta x(s) \rightarrow \mathbf{0}_n$ exponentially fast for all $x(s) \in \bar{\mathcal{A}}$. Then, because $\bar{\mathcal{A}}$ is compact and convex, it follows from Theorem A.4 and Corollary A.1 that all trajectories of $x(s+1) = \mathbf{F}_{\sigma(s)}(x(s))$ converge exponentially fast to a single trajectory, which is denoted by $x^*(s), \forall s \geq 0$. For any $s_1 \in \mathcal{S}$, the value $x^*(s_1)$ depends only on $\gamma_{\sigma(s)}, s < s_1$, i.e. it is independent of $x(0)$. Note that it was established earlier that $x(s) \in \text{int}(\Delta_n), \forall s \geq 0$, which implies $x^*(s) \in \text{int}(\Delta_n), \forall s \geq 0$.

Finally, following the same analysis as in [Jia et al., 2015, pp.393], one can show that $\lim_{s \rightarrow \infty} \zeta(s) = x^*(s)$ and $\lim_{s \rightarrow \infty} \mathbf{W}(x(s)) = \mathbf{X}^*(s) + (\mathbf{I}_n - \mathbf{X}^*(s))\mathbf{C}(s) = \mathbf{W}(x^*(s))$. That is, the dominant left eigenvector $\zeta(s)^\top$ of $\mathbf{W}(s)$ converges to be equal to the unique trajectory $x^*(s)^\top, \forall s \geq 0$, and the influence matrix $\mathbf{W}(x(s))$ converges to an issue-varying matrix function of the unique trajectory $x^*(s)$, with functional form given above. \square

The above result implies that the system Eq. (6.4), with initial conditions satisfying $0 \leq x_i(0) < 1, \forall i$ and $\exists j : x_j(0) > 0$, converges to a trajectory $x^*(s), s \geq 0$ of Eq. (6.4) as $s \rightarrow \infty$. For convenience in future discussions and presentation of results, this will be called the *unique limiting trajectory* of Eq. (6.4). This is a limiting trajectory in the sense that $\lim_{s \rightarrow \infty} [x(s) - x^*(s)] = \mathbf{0}_n$.

Remark 6.2 (Relaxation of the initial conditions). *Theorem 6.1 contains a relaxation of the initial conditions of the original DeGroot–Friedkin model, and provides a more reasonable interpretation from a social context. One can consider $x_i(0)$ as individual i 's estimate of i 's social power (or perceived social power) when the social network is first formed and before discussion begins on issue $s = 0$. The original model requires $x(0) \in \tilde{\Delta}_n$ to avoid an autocratic system (an autocratic system is where $x(s) = \mathbf{e}_i$ for some i , i.e., an individual holds all the social power). However, this is unrealistic because one cannot expect individuals to have estimates such that $\sum_{i=1}^n x_i(0) = 1$. On the other hand, it has been shown that the system trajectory satisfies further, as already commented, $\sum x_i(1) = 1$ and $x(1) \in \text{int}(\Delta_n)$, and then easily $\sum_{i=1}^n x_i(k) = 1, \forall k > 1$ and $x(k) \in \text{int}(\Delta_n)$. Moreover, $x^*(s) \in \text{int}(\Delta_n)$, i.e., $x_i^*(s) > 0, \forall i$ and $\sum_{i=1}^n x_i^*(s) = 1, \forall s$. This mild assumption holds as long as no individual i estimates his/her social power to be autocratic ($x_i(0) = 1$) and at least one individual j has an estimate of his/her own social power that is strictly positive ($\exists j : x_j(0) > 0$).*

The main convergence result is concluded by showing that if the set of switching matrices has a special property (which is sometimes but certainly not always satisfied), then the unique limiting trajectory $x^*(s)$ is in fact a stationary point.

Theorem 6.2. *Suppose that the relative interaction matrix $\mathbf{C}(s)$ of the DeGroot–Friedkin model, satisfying Assumptions 6.1 and 6.2, switches as $\mathbf{C}(s) = \mathbf{C}_{\sigma(s)} \in \mathcal{K}(\tilde{\gamma})$, where*

$\mathcal{K}(\tilde{\gamma}) = \{C_p \in \mathbb{R}^{n \times n} : \gamma_p = \tilde{\gamma}, \forall p \in \mathcal{P} = \{1, 2, \dots, P\}\}$, is the set of C all having the same dominant left eigenvector $\tilde{\gamma}^\top$, with P finite³. Then, the system Eq. (6.4), with initial conditions $0 \leq x_i(0) < 1, \forall i$ and $\exists j : x_j(0) > 0$, converges exponentially fast to a unique point $\mathbf{x}^* \in \text{int}(\Delta_n)$.

There holds $x_i^* < x_j^* \Leftrightarrow \tilde{\gamma}_i < \tilde{\gamma}_j$, for any i, j , where $\tilde{\gamma}_i$ and x_i^* are the i^{th} entry of $\tilde{\gamma}$ and \mathbf{x}^* , respectively. There holds $x_i^* = x_j^* \Leftrightarrow \tilde{\gamma}_i = \tilde{\gamma}_j$.

Proof. The map $F_{\sigma(s)}$ is parametrised simply by the vector $\gamma_{\sigma(s)}$. Under the stated condition of $C(s) = C_{\sigma(s)} \in \mathcal{K}(\tilde{\gamma})$, the map $F_{\sigma(s)}$ is time-invariant. The result in Theorem 5.2 is then used to complete the proof. \square

6.3.1 Extending the Results in Section 5.3

The “contraction-like” property and the subsequent results on an upper bound on the social power at equilibrium and the convergence rate for a class of topologies, i.e. Theorem 5.3, Corollary 5.2 and Lemma 5.2, are now extended to the case of dynamic relative interaction matrices. Naturally, the results are slightly different due to the dynamic topology, e.g. the upper bound result now applies to the entries of the unique limiting trajectory $\mathbf{x}^*(s), s \geq 0$ as opposed to the fixed point \mathbf{x}^* .

Theorem 6.3. Consider the system Eq. (6.4) and suppose that $0 \leq x_i(0) < 1, \forall i \in \mathcal{I}$ and $\exists k \in \mathcal{I} : x_k(0) > 0$. Suppose further that Assumptions 6.1 and 6.2 hold. Define

$$\bar{r}_j = \frac{1 - 2\tilde{\gamma}_j}{1 - \tilde{\gamma}_j}, \quad j \in \{1, \dots, n\} \quad (6.7)$$

where $\tilde{\gamma}_j = \max_{p \in \mathcal{P}} \gamma_{p,j}$, $\mathcal{P} = \{1, \dots, P\}$ with $P < \infty$, and $\gamma_{p,j}$ is the j^{th} entry of the dominant left eigenvector γ_p^\top of C_p . Then, for any $0 < r \leq \bar{r}_j$ and $p \in \mathcal{P}$, there holds

$$x_j = 1 - r \Rightarrow F_{p,j}(\mathbf{x}) < 1 - r \quad (6.8)$$

where $F_{p,j}(\mathbf{x})$ is the j^{th} entry of $F_p(\mathbf{x})$ defined in Eq. (6.5).

Proof. The theorem is proved by checking that, for the given definition of \bar{r}_j , the result in Theorem 5.3 holds separately for every map F_p , $p \in \mathcal{P}$. In other words, for all i, p , $x_i(s) = 1 - r \Rightarrow F_{p,i}(\mathbf{x}(s)) < 1 - r, \forall r \leq \bar{r}_i$. \square

Corollary 6.1 (Upper bound on $x_i^*(s)$). Consider the system Eq. (6.4), and suppose $0 \leq x_i(0) < 1, \forall i \in \mathcal{I}$ and $\exists j \in \mathcal{I} : x_j(0) > 0$. Then, there holds $x_i^*(s) < \tilde{\gamma}_i / (1 - \tilde{\gamma}_i), \forall s$, where $\tilde{\gamma}_j = \max_{p \in \mathcal{P}} \gamma_{p,j}$, $\mathcal{P} = \{1, \dots, P\}$ with $P < \infty$, and $x_i^*(s)$ is the i^{th} entry of the unique limiting trajectory $\mathbf{x}^*(s)$.

Proof. The proof is a straightforward extension of the proof of Corollary 5.2. \square

Lemma 6.1 (Convergence Rate for Dynamic Topology). Let $\mathcal{P} = \{1, \dots, P\}$ with $P < \infty$. For all $p \in \mathcal{P}$, suppose that $C_p \in \mathcal{L}$ where $\mathcal{L} = \{C_p \in \mathbb{R}^{n \times n} : \gamma_{p,i} < 1/3, \forall i\}$ and

³A well-known set is $\mathcal{K}(\mathbf{1}_n/n)$, the set of $n \times n$ doubly-stochastic C .

$\gamma_{p,i}$ is the i^{th} entry of γ_p . Then, there exists a finite s_1 such that, for all $s \geq s_1$, the system Eq. (6.4) contracts to its unique limiting trajectory $\mathbf{x}^*(s)$ with a convergence rate obeying

$$\|\mathbf{x}^*(s) - \mathbf{x}(s+1)\|_1 \leq (2\bar{\beta} - \epsilon)\|\mathbf{x}^*(s) - \mathbf{x}(s)\|_1 \quad (6.9)$$

where $\bar{\beta} = \max_{p \in \mathcal{P}} \max_{i \in \mathcal{I}} \gamma_{p,i} / (1 - \gamma_{p,i}) < 1/2$ and ϵ is a sufficiently small positive constant.

Proof. The proof is an extension of the proof of Lemma 5.2, recalling from the proof of Theorem 6.1 that the Jacobian takes on the same form. \square

Remark 6.3 (Self-Regulation). Exponential convergence to a unique trajectory $\mathbf{x}^*(s)$ can be considered from another point of view as the system Eq. (6.4) **forgetting its initial conditions at an exponential rate**, and is a powerful notion. It implies that sequential discussion of topics combined with reflected self-appraisal is a method of “self-regulation” for social networks, even in the presence of dynamic topology. Consider an individual i who is extremely arrogant, e.g. $x_i(0) = 0.99$. However, individual i is not likeable and others tend to not trust i ’s opinions on any issue, e.g. $c_{ji}(s) \ll 1, \forall j, s$. Then, $\gamma_i(s) \ll 1$ because $\gamma(s)^\top = \gamma(s)^\top \mathbf{C}(s)$ implies $\gamma_i(s) = \sum_{j \neq i} \gamma_j(s) c_{ji}(s)$. According to Corollary 6.1, $x_i^*(s) \leq \bar{\gamma}_i(1 - \bar{\gamma}_i)^{-1}$, and $\bar{\gamma}_i(1 - \bar{\gamma}_i)^{-1}$ is extremely small. Thus, individual i exponentially loses his/her individual social power. An interesting future extension would be to expand on the reflected self-appraisal by modelling individual **behaviour**. For example, one can consider $x_i(s+1) = \phi_i(\zeta_i(s))$ where $\phi_i(\cdot)$ is a smooth function that may capture arrogance or humility of an individual when self-appraising his/her true social power $\zeta_i(s)$.

From the above results, one can also conclude that, for large s , any individual wanting to have an impact on the discussion of topic $s+1$ should focus on ensuring it has a large impact on discussion of the prior topic s . This concept can be applied to Chapter 7.

6.4 Periodically Varying Topology

This section investigates the interesting, special case of periodically varying $\mathbf{C}(s)$ which satisfies Assumption 6.2.

Motivation for Periodic Variations: Consider the example on issue-varying topology in Section 6.1.1 of a government cabinet that meets to discuss the issues of defence, economic growth, social security programs and foreign policy. Since these issues are vital to the smooth running of the country, one expects the issues to be discussed *regularly and repeatedly*. Regular meetings on the same set of issues for decision making/governance/management of a country or company then points to periodically varying $\mathbf{C}(s)$, i.e., social networks with periodic topology.

The system Eq. (6.4), with periodically switching $\mathbf{C}(s)$, can be described by a switching signal $\sigma(s)$ of the form $\sigma(0) = P$, and for $s \geq 1$, $\sigma(Pq + p) = p$,⁴ where $P < \infty$ is the period length, $p \in \mathcal{P} = \{1, 2, \dots, P\}$ and $q \in \mathbb{Z}_{\geq 0}$ is any nonnegative

⁴Any given $s \in \mathcal{S}$ can be uniquely expressed by a given fixed positive integer P , a nonnegative integer q , and positive $p \in \mathcal{P}$, as shown.

integer. Note that in general, $C_i \neq C_j$ for $i, j \in \mathcal{P}$ and $i \neq j$. Theorem 6.1 immediately allows one to conclude that the system Eq. (6.4) with periodic switching converges exponentially fast to the system's unique limiting trajectory $x^*(s)$. This section's key contribution is to obtain additional, insightful information on the limiting trajectory, including an observation that the limiting trajectory itself is periodic.

For simplicity, analysis begins with the assumption that $\mathcal{P} = \{1, 2\}$, i.e., there are two different C matrices, and the switching is of period 2. In the sequel, it will be shown that analysis for $\mathcal{P} = \{1, 2, \dots, P\}$, with arbitrarily large but finite P , is a simple recursive extension on the analysis for $\mathcal{P} = \{1, 2\}$. For the $P = 2$, one obtains

$$x(s+1) = \begin{cases} F_1(x(s)) & \text{if } s \text{ is odd,} \\ F_2(x(s)) & \text{if } s \text{ is even.} \end{cases} \quad (6.10)$$

The periodic system is now transformed into a time-invariant system. Define a new state $z \in \mathbb{R}^{2n}$ (different to the differential transform δz employed in Sections 5.2 and 6.3 for nonlinear contraction analysis) as

$$z(2q) = \begin{bmatrix} z_1(2q) \\ z_2(2q) \end{bmatrix} = \begin{bmatrix} x(2q) \\ x(2q+1) \end{bmatrix}. \quad (6.11)$$

The evolution of $z(2q)$ is studied for $q \in \{0, 1, 2, \dots\}$. Note that

$$z(2(q+1)) = \begin{bmatrix} z_1(2(q+1)) \\ z_2(2(q+1)) \end{bmatrix} = \begin{bmatrix} x(2(q+1)) \\ x(2(q+1)+1) \end{bmatrix}. \quad (6.12)$$

In view of the fact that $x(2(q+1)) = F_1(x(2q+1))$ and $x(2(q+1)+1) = F_2(x(2q+2))$ for any $q \in \{0, 1, 2, \dots\}$, it follows that

$$z(2(q+1)) = \begin{bmatrix} F_1(x(2q+1)) \\ F_2(x(2q+2)) \end{bmatrix}. \quad (6.13)$$

Similarly, notice that $x(2q+1) = F_2(x(2q))$ and $x(2q+2) = F_1(x(2q+1))$ for any $q \in \{0, 1, 2, \dots\}$. From this, for $q \in \{0, 1, 2, \dots\}$, one obtains

$$z(2(q+1)) = \begin{bmatrix} F_1(F_2(z_1(2q))) \\ F_2(F_1(z_2(2q))) \end{bmatrix} = \begin{bmatrix} G_1(z_1(2q)) \\ G_2(z_2(2q)) \end{bmatrix}, \quad (6.14)$$

with the time-invariant composition maps $G_1 = F_1 \circ F_2$ and $G_2 = F_2 \circ F_1$. Thus, one can express the periodic system Eq. (6.10) as the nonlinear time-invariant system

$$z(2q+2) = \bar{G}(z(2q)), \quad \bar{G} = [G_1^\top, G_2^\top]^\top \quad (6.15)$$

Theorem 6.4. *There exists a unique periodic sequence $x^*(s)$ for the system Eq. (6.10), with*

map F_p given in Eq. (6.5) for $p = 1, 2$, obeying

$$\mathbf{x}^*(s) = \begin{cases} \mathbf{z}_1^* & \text{if } s \text{ is odd,} \\ \mathbf{z}_2^* & \text{if } s \text{ is even,} \end{cases} \quad (6.16)$$

where $\mathbf{z}_1^* \in \text{int}(\Delta_n)$ and $\mathbf{z}_2^* \in \text{int}(\Delta_n)$ are the unique fixed points of, respectively, \mathbf{G}_1 and \mathbf{G}_2 , defined above Eq. (6.15). For all initial conditions $0 \leq x_i(0) < 1, \forall i \in \mathcal{I}$ and $\exists j \in \mathcal{I} : x_j(0) > 0$, $\lim_{s \rightarrow \infty} [\mathbf{x}(s) - \mathbf{x}^*(s)] = \mathbf{0}_n$ exponentially fast.

Proof. As mentioned above, one can immediately apply Theorem 6.1 to show that $\lim_{s \rightarrow \infty} [\mathbf{x}(s) - \mathbf{x}^*(s)] = \mathbf{0}_n$. This proof focuses on using the time-invariant transformation to show that $\mathbf{x}^*(s)$ has the properties described in the theorem statement.

Part 1: In this part, the map $\mathbf{G}_i, i = 1, 2$ is proved to have at least one fixed point. First, note that it was proved in Theorem 6.1 that the system Eq. (6.4), with initial conditions $0 \leq x_i(0) < 1, \forall i$ and for at least one $j, x_j(0) > 0$, will have $\mathbf{x}(s) \in \text{int}(\Delta_n)$ for all $s > 0$, which implies that $\mathbf{x}^*(s) \in \text{int}(\Delta_n)$. For $p = 1, 2$, Property 6.1 established that F_p is continuous on Δ_n . The composition of continuous functions is continuous, which implies $\mathbf{G}_1 = \mathbf{F}_1 \circ \mathbf{F}_2 : \Delta_n \mapsto \Delta_n$ and $\mathbf{G}_2 = \mathbf{F}_2 \circ \mathbf{F}_1 : \Delta_n \mapsto \Delta_n$ are continuous.

The proof of Theorem 6.1 also showed that for all $p, F_p \in \bar{\mathcal{A}}$ where $\bar{\mathcal{A}} = \{\mathbf{x} \in \Delta_n : 1 - \bar{r} \geq x_i \geq 0, \forall i \in \mathcal{I}\}$ and \bar{r} is some small strictly positive constant. For the system in Eq. (6.10) with $p = 1, 2$, it follows that $\mathbf{F}_1(\bar{\mathcal{A}}) \subset \bar{\mathcal{A}} \Rightarrow \mathbf{F}_2(\mathbf{F}_1(\bar{\mathcal{A}})) \subset \bar{\mathcal{A}}$, which implies that $\mathbf{G}_1(\bar{\mathcal{A}}) \subset \bar{\mathcal{A}}$. Similarly, $\mathbf{G}_2(\bar{\mathcal{A}}) \subset \bar{\mathcal{A}}$. Brouwer's Fixed Point Theorem then implies that there exists at least one fixed point $\mathbf{z}_1^* \in \bar{\mathcal{A}}$ such that $\mathbf{z}_1^* = \mathbf{G}_1(\mathbf{z}_1^*)$ (respectively at least one fixed point $\mathbf{z}_2^* \in \bar{\mathcal{A}}$ such that $\mathbf{z}_2^* = \mathbf{G}_2(\mathbf{z}_2^*)$) because \mathbf{G}_1 (respectively \mathbf{G}_2) is a continuous function on the compact, convex set $\bar{\mathcal{A}}$.

Part 2: In this part, it is proved that the unique limiting trajectory of Eq. (6.10) obeys Eq. (6.16) and that \mathbf{z}_1^* and \mathbf{z}_2^* are the unique fixed point of \mathbf{G}_1 and \mathbf{G}_2 , respectively. Let \mathbf{z}_1^* be a fixed point of \mathbf{G}_1 . Observe that $\mathbf{z}_1^* = \mathbf{F}_2(\mathbf{F}_1(\mathbf{z}_1^*))$. Define $\mathbf{z}_2^* = \mathbf{F}_1(\mathbf{z}_1^*)$. One has that $\mathbf{z}_1^* = \mathbf{F}_2(\mathbf{z}_2^*)$. Observe that $\mathbf{F}_1(\mathbf{z}_1^*) = \mathbf{F}_1(\mathbf{F}_2(\mathbf{z}_2^*))$, which implies that $\mathbf{z}_2^* = \mathbf{F}_1(\mathbf{F}_2(\mathbf{z}_2^*)) = \mathbf{G}_2(\mathbf{z}_2^*)$. In other words, \mathbf{z}_2^* is a fixed point of \mathbf{G}_2 (but at this stage, and similarly for \mathbf{z}_1^* , its uniqueness has not yet been proved).

Next, it will be established that the unique limiting trajectory is the periodic sequence in Eq. (6.16). Recall that Theorem 6.1 yields the conclusion that *all trajectories* of Eq. (6.10) converge exponentially fast to a unique limiting trajectory $\mathbf{x}^*(s) \in \text{int}(\Delta_n)$. It follows, from Eq. (6.15) and the definition of $\mathbf{z}(2q)$, that for all $s \geq 0$, Eq. (6.16) is a trajectory of the system Eq. (6.10); the critical point here is that Eq. (6.16) holds for all s . Combining these arguments, it is clear that Eq. (6.16) is precisely the unique limiting trajectory.

Last, it will be shown that \mathbf{z}_1^* and \mathbf{z}_2^* are the unique fixed point of \mathbf{G}_1 and \mathbf{G}_2 , respectively. To this end, suppose that, to the contrary, at least one of \mathbf{z}_1^* and \mathbf{z}_2^* is not unique. Without loss of generality, suppose in particular that $\mathbf{z}_1' \neq \mathbf{z}_1^*$ is any other

fixed point of G_1 . Then, $z'_2 = F_1(z'_1)$ is a fixed point of G_2 , and

$$x(s) = \begin{cases} z'_1 & \text{if } s \text{ is odd,} \\ z'_2 & \text{if } s \text{ is even,} \end{cases} \quad (6.17)$$

is a trajectory of Eq. (6.10) that holds for all $s \geq 0$, and is *different from the trajectory Eq. (6.16) because $z'_1 \neq z_1^*$* . On the other hand, Theorem 6.1 implies that all trajectories of Eq. (6.10) converge exponentially fast to a *unique limiting trajectory*, which yields a contradiction. Thus, z_1^* and z_2^* are the unique fixed point of G_1 and G_2 , respectively, and the system in Eq. (6.10) converges exponentially fast to the unique limiting trajectory Eq. (6.16), which completes the proof⁵. \square

The generalisation is now provided for periodically switching topology $C(s) = C_{\sigma(s)}$, where $\sigma(s)$ is of the form $\sigma(0) = P$, and for $s \geq 1$, $\sigma(Pq + p) = p$. Here, $2 \leq P < \infty$, $p \in \mathcal{P} = \{1, 2, \dots, P\}$ and $q \in \mathbb{Z}_{\geq 0}$. The periodic DeGroot–Friedkin model is described by

$$x(s+1) = \begin{cases} F_P(x(s)) & \text{for } s = 0 \\ F_p(x(s = Pq + p)) & \text{for all } s \geq 1. \end{cases} \quad (6.18)$$

A transformation of Eq. (6.18) to a time-invariant system follows a procedure similar to the one detailed for $P = 2$ (in the signal processing and control literature, this technique is sometimes referred to as “lifting”). A new state variable $z \in \mathbb{R}^{Pn}$ is defined as

$$z(Pq) = \begin{bmatrix} z_1(Pq) \\ z_2(Pq) \\ \vdots \\ z_P(Pq) \end{bmatrix} = \begin{bmatrix} x(Pq) \\ x(Pq+1) \\ \vdots \\ x(Pq+P-1) \end{bmatrix}, \quad (6.19)$$

and the evolution of $y(Pq)$ is studied for $q \in \{0, 1, \dots\}$. It follows that

$$z_p(P(q+1)) = x(P(q+1) + p - 1), \quad \forall p \in \mathcal{P}.$$

Following the logic in the $P = 2$ case, but with the precise steps omitted, one obtains

$$z(P(q+1)) = \begin{bmatrix} F_{P-1}(F_{P-2}(\dots(F_P(z_1(Pq)))))) \\ F_P(F_{P-1}(\dots(F_1(z_2(Pq)))))) \\ \vdots \\ F_{P-2}(F_{P-1}(\dots(F_P(z_P(Pq)))))) \end{bmatrix} = \bar{G}(z(Pq)), \quad (6.20)$$

⁵During the thesis examination process, an examiner has identified an alternative proof, which we summarise here for the interested reader. The existence of $x^*(s)$ was established in Theorem 6.1. Since $x(2s+2)$ is also a trajectory, $\|x^*(2s+2) - x^*(2s)\|$ converges to 0 exponentially fast. From this, one has that for any $s \geq 0$ and $k \in \mathbb{N}$, $\|x^*(s+2k) - x(2s)\|$ converges to 0 exponentially fast; that is $x^*(2s)$ is a Cauchy sequence in $\bar{\mathcal{A}}$, which must converge to a point, denoted as z_1^* . From Eq. (6.14), and by continuity, z_1^* must be a fixed point of $F_1(F_2(\cdot))$. A similar approach shows that $x^*(2s+1)$ must converge to a point, z_2^* , which must be a fixed point of $F_2(F_1(\cdot))$.

where $\bar{G}(z) = [G_1(z_1)^\top, G_2(z_2)^\top, \dots, G_P(z_P)^\top]^\top$. This leads to the following generalisation of Theorem 6.4.

Theorem 6.5. *There exists a unique periodic sequence $x^*(s)$ for the system Eq. (6.18), with map F_p given in Eq. (6.5) for $p = 1, 2, \dots, P$, which for any nonnegative integer q , obeys*

$$x^*(Pq + p - 1) = z_p^*, \text{ for all } p \in \{1, 2, \dots, P\} \quad (6.21)$$

where $z_p^* \in \text{int}(\Delta_n)$ is the unique fixed point of G_p defined in Eq. (6.20). For all initial conditions satisfying $0 \leq x_i(0) < 1, \forall i \in \mathcal{I}$ and for at least one $j \in \mathcal{I}$, $x_j(0) > 0$, $\lim_{s \rightarrow \infty} [x(s) - x^*(s)] = \mathbf{0}_n$ exponentially fast.

Proof. The proof is obtained by recursively applying the same techniques used in the proof of Theorem 6.4. The detailed mechanical calculations are omitted. \square

This section is concluded by noting that Theorem 6.3, Lemma 6.1 and Corollary 6.1 are all applicable to the periodic system Eq. (6.18) because Eq. (6.18) is just a special case of the general switching system Eq. (6.4).

6.5 Simulations

In this section, simulations are provided to illustrate the key conclusions of this chapter. To begin, a simulation for general switching topology is given, and then the special case of periodically-varying topology is illustrated. The set of topologies is given as $\mathcal{C} = \{C_1, \dots, C_5\}$, i.e., $\mathcal{P} = \{1, 2, \dots, 5\}$. The switching signal $\sigma(s)$ is generated such that for any given s , there is equal probability that $\sigma(s) = p, \forall p \in \mathcal{P}$. The precise numerical forms of C_p are stated in Section 6.7.

Figure 6.1 shows the evolution of individual social power over a sequence of issues for the system as described in the above paragraph, initialised from an arbitrarily chosen initial condition vector⁶, $\hat{x}(0) = [0.95, 0.95, 0.95, 0, 0, 0]^\top$. For each individual, with $\bar{\gamma}_i = \max_{p \in \mathcal{P}} \gamma_{p,i}$, it is computed that $\bar{\gamma}_1 = 0.474, \bar{\gamma}_2 = 0.237, \bar{\gamma}_3 = 0.244, \bar{\gamma}_4 = 0.244, \bar{\gamma}_5 = 0.244, \bar{\gamma}_6 = 0.239$. Note that $\sum_i \bar{\gamma}_i \neq 1$ in general due to the definition of $\bar{\gamma}_i$. Corollary 6.1 yields that $x^*(s) \leq [0.9, 0.311, 0.323, 0.323, 0.323, 0.314]$, and this can be seen in Fig. 6.1. Since only $\bar{\gamma}_1 > 1/3$, it is observed that after the first 10 or so issues, only $x_1^*(s) > 0.5$. That is, in the limit, only individual 1 may have more than half the social power at some issue. Note that $x_4^*(s) > 0$ for all s , although for several issues, $x_4^*(s)$ is close to 0.

Figure 6.2 shows the system with a different arbitrarily selected vector of initial conditions $\tilde{x}(0) = [0.1, 0.325, 0, 0.8, 0.45, 0.7]^\top \neq \hat{x}(0)$. Notice that for the initial condition vector $\hat{x}(0)$ individuals 1, 2, 3 have large perceived social power $\hat{x}_i(0) = 0.95$, while individuals 4, 5, 6 have $\hat{x}_i(0) = 0$. In contrast, for the other initial condition vector $\tilde{x}(0)$, the perceived social power $\tilde{x}_i(0)$ is large for $i = 4, 6$ and small for $i = 1, 2, 3$. Through sequential discussion and reflected self-appraisal, the initial conditions are

⁶The initial condition vector $\hat{x}(0)$ also illustrates the relaxed initial condition assumption.

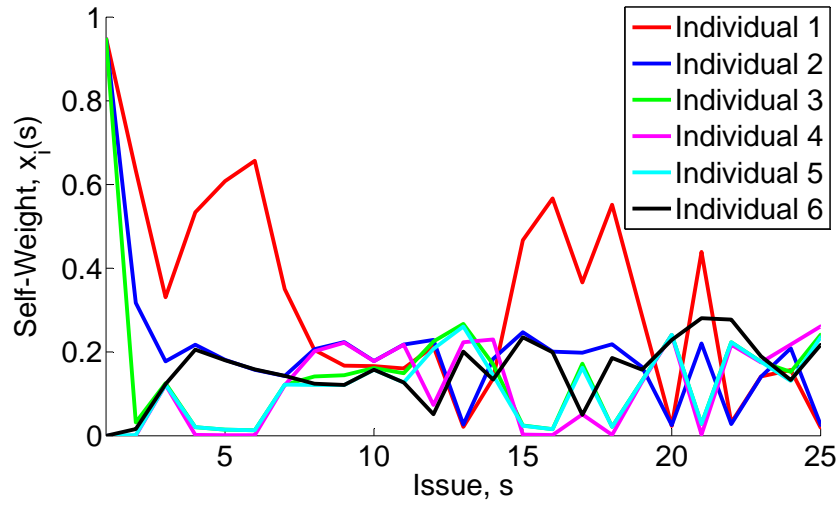


Figure 6.1: Evolution of individuals' social powers $x(s)$ for initial condition vector $\hat{x}(0)$. Viewed in conjunction with Fig. 6.2, it is clear that each individual i 's social power trajectory $x_i(s)$ converges to $x_i^*(s)$ by about $s = 10$.

exponentially forgotten and both Fig. 6.1 and 6.2 show convergence to the same unique limiting trajectory $x^*(s)$ by about $s = 10$. This is highlighted in Fig. 6.3, which displays the individual social powers of selected individuals 1, 3 and 6. The solid and dotted lines correspond to initial condition vectors $\tilde{x}(0)$ and $\hat{x}(0)$, respectively.

Simulations for periodically-varying topology are now presented. The period is selected to be $P = 4$, with $C_3 = C_4$. That is, the switching signal is $\sigma(s) = [3, 1, 2, 3, 3, 1, 2, 3, 3, 1, 2, \dots]$. The matrices used are C_1, C_2, C_3 given in Section 6.7. The same two initial condition vectors $\hat{x}(0)$ and $\tilde{x}(0)$ from the previous simulation are used. The simulation result for $\hat{x}(0)$ is shown in Fig. 6.4. Fig. 6.5 shows a comparison of the social power trajectory for select individuals 1, 3 and 6, with the two different vectors of initial conditions $\tilde{x}(0)$ and $\hat{x}(0)$. It can clearly be seen that the initial conditions are forgotten exponentially fast, and the unique limiting trajectory $x^*(s)$ is a periodic sequence.

Remark 6.4. It is worth noting that the unique limiting trajectory depends on the *sequence of switching topologies*, i.e. the switching signal $\sigma(s)$. For example, given the same C_p in the periodically-varying topology simulation, but with a different switching signal (i.e. ordering of $C(s)$), the unique limiting trajectory will also be different.

6.6 Conclusions

This chapter has extended the DeGroot–Friedkin model through the incorporation of issue-varying, dynamic relative interpersonal interactions. The nonlinear contraction analysis framework of Chapter 5 has been used to study the system with dynamic

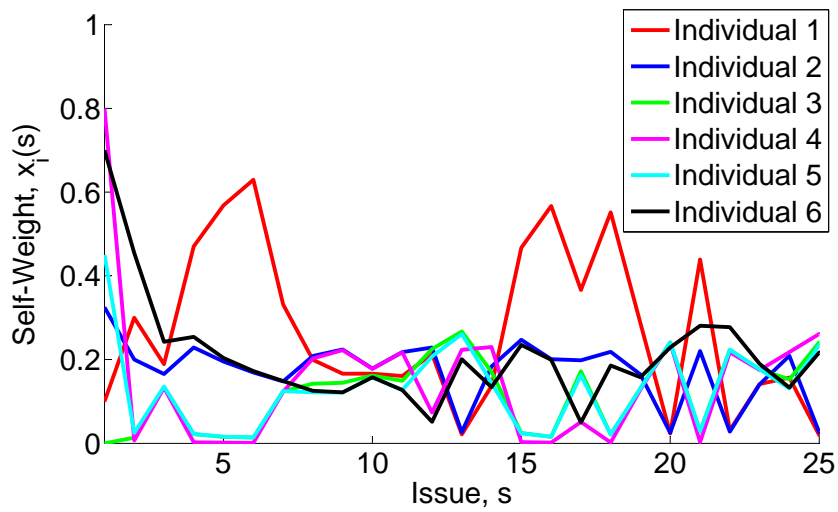


Figure 6.2: Evolution of individuals' social powers $x_i(s)$ for initial condition vector $\tilde{x}(0)$. Viewed in conjunction with Fig. 6.1, it is clear that each individual i 's social power trajectory $x_i(s)$ converges to $x_i^*(s)$ by about $s = 10$.

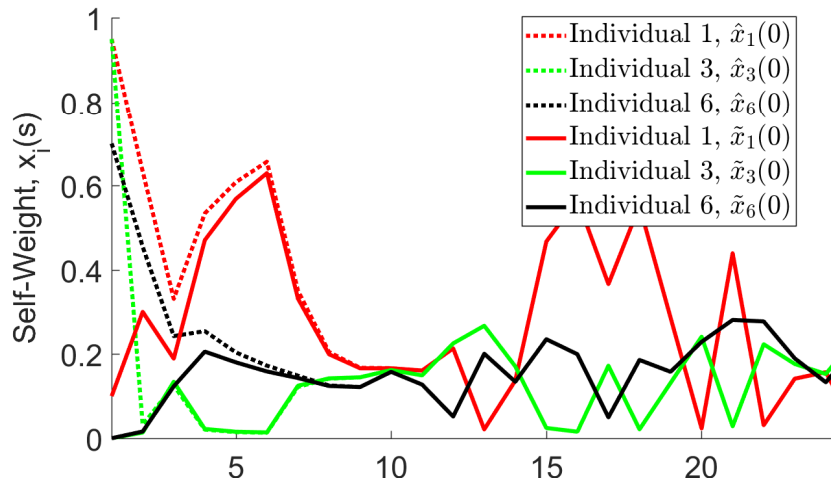


Figure 6.3: Evolution of selected individuals' social powers $x_i(s)$, for $i = 1, 3, 6$. The trajectory $x_i(s)$, for $i = 1, 3, 6$, beginning from the two different initial conditions, converges to the same trajectory by about $s = 10$, which is precisely the unique limiting trajectory $x_i^*(s)$.

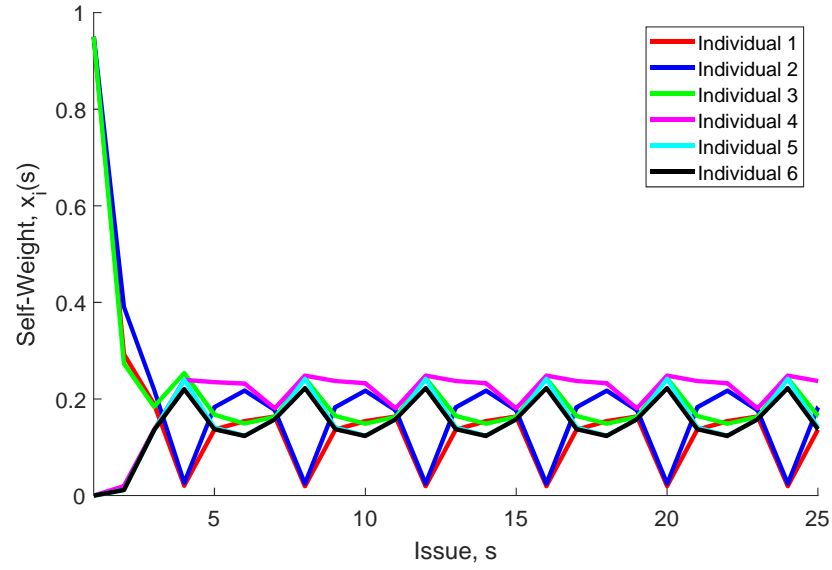


Figure 6.4: Evolution of individuals' social powers $x(s)$ for initial condition vector $\hat{x}(0)$, with periodically-varying topology. Clearly, $x(s)$ becomes a periodic sequence by about $s = 10$.

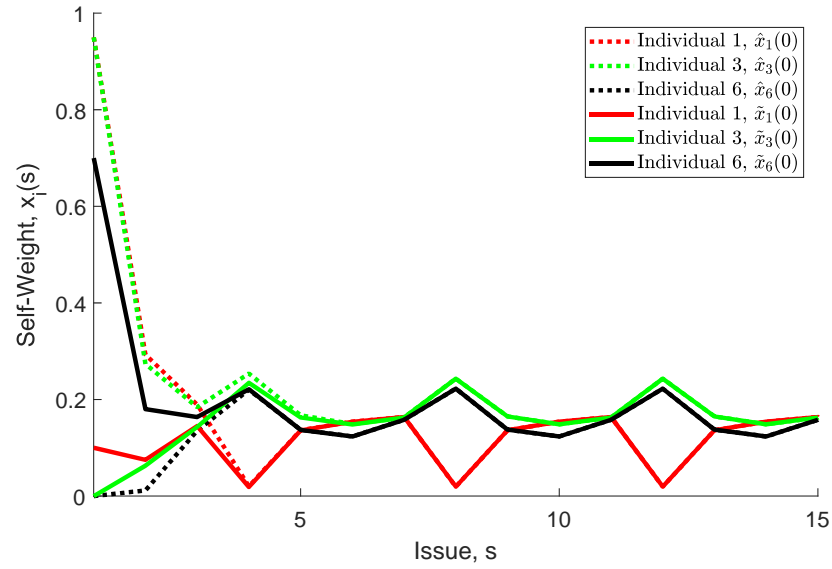


Figure 6.5: Evolution of selected individuals' social powers $x_i(s)$: a comparison of different initial condition vectors $\hat{x}(0)$ and $\tilde{x}(0)$, with periodically-varying topology. The trajectory $x_i(s)$, for $i = 1, 3, 6$, beginning from the two different initial conditions, converges to the same periodic sequence $x_i^*(s)$ by about $s = 10$.

topology, leading to the establishment of a general exponential convergence result. This has been interpreted from the context of a social network discussing a sequence of topics, where the DeGroot opinion process is coupled with the socio-psychological mechanism of reflected self-appraisal. It is concluded that this results in every individual forgetting his/her initial (perceived) social power exponentially fast. In the limit of the topic sequence, the dynamics of each individual's social power depends only on the dynamically-varying topology associated with the topic sequence. An upper bound on an individual's limiting social power trajectory and the convergence rate for a class of topologies has also been obtained. Last, the special case of periodically switching topologies has been investigated in further detail.

6.7 Appendix: Simulation Parameters

The C used in the simulation are arbitrarily chosen by the author to illustrate the results of this chapter. The matrices were not obtained from real-world networks.

$$\begin{aligned} C_1 &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, & C_2 &= \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0.8 & 0 & 0 & 0 & 0 & 0.2 \\ 0 & 0.1 & 0 & 0 & 0 & 0.9 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \\ C_3 &= \begin{bmatrix} 0 & 0 & 0 & 0.2 & 0 & 0.8 \\ 0.3 & 0 & 0.7 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0.5 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0.75 & 0 & 0 & 0.25 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, & C_4 &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0.85 & 0.15 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0.3 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0.9 & 0 & 0 & 0.1 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \\ C_5 &= \begin{bmatrix} 0 & 0.5 & 0 & 0 & 0 & 0.5 \\ 0.9 & 0 & 0.1 & 0 & 0 & 0 \\ 0.9 & 0 & 0 & 0 & 0 & 0.1 \\ 0.9 & 0.1 & 0 & 0 & 0 & 0 \\ 0.9 & 0 & 0 & 0.1 & 0 & 0 \\ 0.9 & 0 & 0 & 0 & 0.1 & 0 \end{bmatrix}. \end{aligned}$$

Modification of Social Dominance in Autocratic Networks

Chapter Summary

The previous two chapters focused on establishing the properties of dynamics of the individuals' social powers, including convergence results. This chapter returns to the original DeGroot–Friedkin model with constant relative interaction topology, and considers a network modification problem. In particular, the star topology is studied, because this special type of topology leads to a single individual (the centre node of the star) accumulating all of the social power. Different strategies, via the insertion of new individuals and/or new interpersonal relationships, are proposed to modify the star topology so that the centre individual no longer has the largest social power. Necessary and sufficient conditions on the strength of the new interpersonal relationships are obtained for each strategy. Interpretations are provided in the context of social networks, which leads to several insightful conclusions.

7.1 Introduction

In the study of dynamic complex networks (within which social networks are a sub-field), one may focus on the temporal evolution of certain variables of interest (e.g. opinions in influence networks) and on establishing graphical conditions or conditions on certain parameters associated with the nodes, which guarantee the variables converge to some (perhaps desired) limiting configuration. Examples of such works can be found in parts of Chapters 3, 5 and 6. For the DeGroot model, a strongly connected and aperiodic influence network is sufficient for ensuring that opinions converge to a consensus (see Section 2.3.1). Once convergence of the variables of interest to a limiting configuration has been established, one may then be interested in investigating how aspects of the network topology *affect or determine* the limiting configuration. For example, in the DeGroot model, a certain left eigenvector of the influence matrix W (along with the initial conditions) determines the final consensus value, while the i^{th} entry of this particular eigenvector can be considered as the relative “impact” of individual i in the opinion discussion (see Remark 5.1 and

Section 2.3.1). Logically, it follows that a topic of considerable interest is in the modification of the network topology to *maximise a desirable* (or minimise an undesirable) property, e.g. the relative “impact” of a specific individual.

The number of works which consider network modifications are vast, and the approach and objective depends on the context of the network and the dynamical process being modelled (financial, epidemic, social or power grid); only a small selection is presented here. In [Kempe et al., 2003], the most influential individual in spreading ideas on a social network was identified. The optimal edge modification of a network to minimise the spreading of a virus was studied in [Preciado et al., 2014; Zhai et al., 2013]. [Wang et al., 2016] studied optimal node addition while maintaining structural controllability. Systematic modification of the network structure to improve performance with respect to an objective function, e.g. reducing the expected number of attacks arising from a terrorist network, was studied in [Mellon et al., 2016]. Strategies for adding edges to reduce the influence of a new individual in the DeGroot model framework were studied in [Amelkin and Singh, 2017], while PageRank optimisation has been investigated in [Fercoq et al., 2013; Csáji et al., 2014].

According to French Jr. and Snyder, “*leadership* is the potential social influence of one part of the group over another” [French Jr and Snyder, 1959]. From the perspective of opinion dynamics, a leader can therefore be seen as an individual or a group of individuals that has a disproportionate amount of control over the opinion discussion process. In the context of social influence, one can therefore refer to a leader/leader group as the *socially dominant* individual/group of individuals in the network. The fact that social influence tends to accumulate with one individual or a subgroup of individuals in a social network is reported empirically in [Friedkin, 2011] and theoretically in [Jia et al., 2015]. This individual or subgroup is defined explicitly by the interpersonal relationships in the network topology. Motivated by this concept of social dominance/leadership, and using the DeGroot–Friedkin model to describe the social power evolution in the network, this chapter considers modifications to the network topology to cause social dominance to shift from one individual to another. In particular, this chapter focuses on modifications to star graphs¹, an example of which is given in Fig. 7.1. Using this example figure for discussion, it was found in [Jia et al., 2015] that for almost all initial values of the individuals’ social powers, all of the social power accumulates with the centre individual 1. It was proved in Chapter 5, Corollary 5.1 that for all non-star graphs, the equilibria in which one individual holds all of the social power are unstable. Also, for a star graph, the equilibria where an individual other than the centre individual holds all of the social power are unstable. In other words, the star graph is the only topology type in which an *autocratic configuration* (where a single individual holds all of the social power) occurs for generic initial conditions. This provides the key motivation to investigate modifications on the star graph.

In particular, this chapter will consider a number of basic modifications to the star graph, which may be generalised in the future. Each modification involves the intro-

¹As will be formally defined, a graph is said to have star topology if there exists a node v_i , called the centre node, such that every edge of the graph is either to or from v_i .

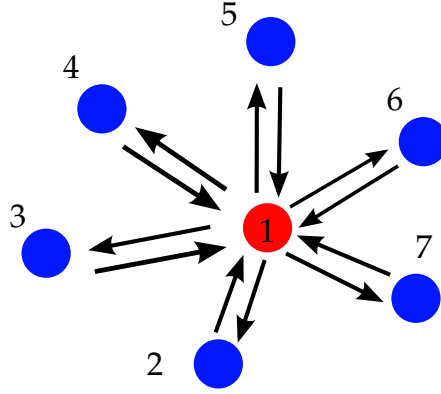


Figure 7.1: Star Topology with red centre node v_1 , with $n = 7$ total individuals.

duction of new individuals (nodes) and/or interpersonal interactions (edges) into the star graph. Existing results indicate that the centre individual will no longer hold all of the social power. However, if the perturbations away from a star topology are small (e.g. the new edge weights are close to 0) then the centre individual will continue to have a disproportionate amount of social power. This chapter derives necessary and sufficient conditions in the form of explicit inequalities involving interpersonal interaction strengths that ensure the centre individual is no longer the socially dominant individual. I.e., conditions guaranteeing that some other identifiable individual will have greater social power at equilibrium than the centre individual are obtained. The results for each modification are discussed in its social context, which yields in each case intuitive interpretation. In one case, it is identified that all new nodes should attach onto the same non-star node in the star graph, rather than attach to different non-star nodes. In other cases, it is identified that certain non-centre nodes need to cooperate, by having all such nodes increase certain associated edge weights, in order to change the social dominance.

7.1.1 Chapter Organization

The chapter is organised as follows. Section 7.2 defines the problem set-up and the topology modifications to be studied. Section 7.3 presents conditions for each modification, and discusses their social implications. Illustrative simulations are given in Section 7.4, with conclusions in Section 7.5. Proofs are given in Appendix Section 6.7.

7.2 Problem Set-up and Proposed Modification Strategies

A special topology in the context of the DeGroot–Friedkin model is the “star topology”, whose definition was given in Definition 5.1, and repeated below for the convenience of the reader. The topology is special because, for $n \geq 3$ individuals and generic initial condition, the set of all star graphs is the *only set of network topologies* for which *the individual at the centre of the star holds all of the social power* in

the limit of the issue sequence. This chapter investigates *how additional nodes and/or edges strategically connected to a star topology can change the social power at equilibrium, \mathbf{x}^** , and in particular to ensure that the centre individual no longer has the greatest social power. To that end, definitions and existing results are now provided to aid in describing the problem and the discussing of new results. To begin, an assumption on the network topology and the DeGroot–Friedkin model dynamics are restated. Note that the term “node” will be used interchangeably with “individual”.

Assumption 7.1. *The graph $\mathcal{G}[\mathbf{C}]$, with $n \geq 3$ nodes, is strongly connected and no node v_i has a self-loop. Moreover, the relative interaction matrix \mathbf{C} is row-stochastic.*

The vector of social powers $\mathbf{x}(s) = [x_1(s), \dots, x_n(s)]^\top$ evolves as

$$\mathbf{x}(s+1) = \mathbf{F}(\mathbf{x}(s)), \quad (7.1)$$

where the nonlinear map $\mathbf{F}(\mathbf{x})$ is defined as

$$\mathbf{F}(\mathbf{x}) = \begin{cases} \mathbf{e}_i & \text{if } \mathbf{x} = \mathbf{e}_i \text{ for any } i \\ \alpha(\mathbf{x}) \begin{bmatrix} \frac{\gamma_1}{1-x_1} \\ \vdots \\ \frac{\gamma_n}{1-x_n} \end{bmatrix} & \text{otherwise} \end{cases}, \quad (7.2)$$

with $\alpha(\mathbf{x}) = 1 / \sum_{i=1}^n \frac{\gamma_i}{1-x_i}$, and $\boldsymbol{\gamma}^\top = [\gamma_1, \dots, \gamma_n]$ is the dominant left eigenvector of \mathbf{C} , and \mathbf{e}_i is the i^{th} canonical unit vector (see Section 2.1). For convenience, the following definitions are restated. The n -simplex is $\Delta_n = \{\mathbf{x} \in \mathbb{R}^n : 0 \leq x_i, \mathbf{1}_n^\top \mathbf{x} = 1\}$, and define $\tilde{\Delta}_n = \Delta_n \setminus \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ and $\text{int}(\Delta_n) = \{\mathbf{x} \in \mathbb{R}^n : 0 < x_i, \mathbf{1}_n^\top \mathbf{x} = 1\}$ as the simplex excluding the corner points and the interior of the simplex, respectively.

Next, the definition of a star topology, and two convergence results (originally from [Jia et al., 2015] and adjusted to incorporate new results presented Theorems 5.2 and 6.1, and Corollary 5.1) are stated

Definition 7.1 (Star topology). *A strongly connected graph $\mathcal{G}[\mathbf{C}]$ has star topology if \exists a node v_i , called the centre node, such that every edge of $\mathcal{G}[\mathbf{C}]$ is either to or from v_i .*

Lemma 7.1. *Suppose that $n \geq 3$, and suppose further that \mathcal{G} has star topology, which without loss of generality has centre node v_1 . Let $\mathcal{G}[\mathbf{C}]$ satisfy Assumption 7.1. Then, for the system Eq. (7.1) with initial condition satisfying $0 \leq x_i(0) < 1, \forall i \in \mathcal{I}$ and for at least one $j \in \mathcal{I}, x_j(0) > 0$, $\lim_{s \rightarrow \infty} \mathbf{x}(s) = \mathbf{e}_1$ asymptotically.*

Theorem 7.1. *For $n \geq 3$, consider the DeGroot–Friedkin dynamical system Eq. (7.1) with $\mathcal{G}[\mathbf{C}]$ satisfying Assumption 7.1. Assume further that the graph $\mathcal{G}[\mathbf{C}]$ does not have star topology. Then,*

- (i) *For all initial conditions satisfying $0 \leq x_i(0) < 1, \forall i \in \mathcal{I}$ and for at least one $j \in \mathcal{I}, x_j(0) > 0$, the self-weight vector $\mathbf{x}(s)$ converges to \mathbf{x}^* exponentially fast as $s \rightarrow \infty$, where $\mathbf{x}^* \in \text{int}(\Delta_n)$ is the unique fixed point satisfying $\mathbf{x}^* = \mathbf{F}(\mathbf{x}^*)$.*

- (ii) There holds $x_i^* < x_j^*$ if and only if $\gamma_i < \gamma_j$ for any i, j , where γ_i is the i^{th} entry of the dominant left eigenvector γ^\top . There holds $x_i^* = x_j^*$ if and only if $\gamma_i = \gamma_j$.
- (iii) The unique fixed point \mathbf{x}^* is determined only by γ , and is independent of $\mathbf{x}(0)$.
- (iv) The fixed points \mathbf{e}_i , for all $i \in \mathcal{I}$, are unstable equilibria of Eq. (7.1).

The following definitions of an autocratic network and a socially dominant individual are introduced to aid future discussions.

Definition 7.2 (Autocratic Network). A social network is said to be an autocratic configuration, with node v_i being the autocrat, if $\mathbf{x}(s) = \mathbf{e}_i$, for some $s \geq 0$.

Definition 7.3 (Social dominance/leadership). Node v_i is said to be the socially dominant/leader node in the network if $x_i^* > x_j^*$ for all $j \neq i$. In other words, at equilibrium, the social power of individual i is greater than the social power of any other individual $j \neq i$.

Remark 7.1 (Autocratic tendency). Lemma 7.1 has an important social connotation. As introduced in previous chapters, one can consider $x_i(0)$ as individual i 's *estimate* of i 's social power when the social network is first formed, before any discussions occur. For any initial estimate where no individual i believes $x_i(0) = 1$ and at least one individual j has $x_j(0) > 0$, the star topology network with centre node v_1 tends to the autocratic configuration $\mathbf{x}^* = \mathbf{e}_1$. Recall from the proof of Theorem 6.1 that $\lim_{s \rightarrow \infty} \zeta(s) = \mathbf{x}^*$, where $\zeta^\top(s)$ is the dominant left eigenvector of the influence matrix $\mathbf{W}(s)$ in the DeGroot opinion dynamics for issue s , $\mathbf{y}(s, t+1) = \mathbf{W}(s)\mathbf{y}(s, t)$ (see Section 5.1.1). Moreover, for any issue s , one has that $\lim_{t \rightarrow \infty} \mathbf{y}(s, t) = (\zeta^\top(s)\mathbf{y}(s, 0)) \mathbf{1}_n$. This implies that, for initial issues, opinion discussion will occur with everyone contributing to the final consensus value. However, the centre individual increasingly guides the outcome of discussions until, for $s = \infty$, there holds $\zeta(s) = \mathbf{x}^* = \mathbf{e}_1$, which implies that only the centre individual influences the final consensus value for issue s .

The above material motivates the problem investigated in this chapter, which can be considered as being embedded in the following objective

Objective 7.1. For a star graph $\mathcal{G}[\mathbf{C}]$ satisfying Assumption 7.1, investigate modifications to the graph $\mathcal{G}[\mathbf{C}]$ via the introduction of new nodes and/or new edges. In particular, given a modified graph, conditions on the strength of edge weights c_{ij} should be identified which guarantee that a selected individual v_k has social dominance.

Remark 7.2. In [Jia et al., 2015], the constant entries c_{ij} of \mathbf{C} are termed “relative interpersonal weights”, a term which is used throughout Part II. However, one can also consider c_{ij} as the amount of relative “trust” individual i allocates to individual j . In other words, c_{ij} captures the strength of a unidirectional interpersonal relationship.

This chapter focuses on comparing the social power of individuals within the network at equilibrium, i.e. when $s \rightarrow \infty$. The equilibrium value x_i^* will therefore be referred to as the social power of individual i when there is no ambiguity (as opposed to the evolving $x_i(s)$ when $s < \infty$). In light of the above discussion, and in

particular the social interpretation of Lemma 7.1 given in Remark 7.1, it is evident that there is interest in studying star graphs and modifications to the star topology. In particular, modifications are considered that turn a star $\bar{\mathcal{G}}[\bar{\mathcal{C}}]$ into a non-star $\mathcal{G}[\mathcal{C}]$ via the introduction of *new* nodes (individuals) or edges (interpersonal relationships). From Lemma 7.1 and Theorem 7.1, it immediately follows that $x_i^* < 1$ for the original centre node v_i , but if the perturbation from star topology (caused by the new edges) is small, one expects that v_i remains *socially dominant*. What will be shown is that if the interpersonal weights associated with these new edges exceed a given threshold, the socially dominant node changes from v_i to some other identifiable node.

To aid in the discussion and presentation of results, some node types are defined. For a given star graph $\bar{\mathcal{G}}[\bar{\mathcal{C}}]$, with centre node v_1 , all the other nodes are called *subject nodes* in the sense that they are subjects to the autocrat centre node. As already noted, this chapter will study how the autocracy of node v_1 can be disrupted; this motivates the definition of a new type of node. An *attacker node* is a node v_j not in the original star, which forms edges e_{ji}, e_{ij} with a subject node v_i . In doing so, the graph $\bar{\mathcal{G}}[\bar{\mathcal{C}}]$ is modified to become $\mathcal{G}[\mathcal{C}]$, which is no longer a star. As it will become apparent, the weights c_{ji} and c_{ij} can modify/reduce the social power x_1^* of the centre node v_1 . In other words, v_j *attacks* the social dominance of v_1 . Note that two edges, e_{ji}, e_{ij} must be formed to ensure that $\mathcal{G}[\mathcal{C}]$ remains strongly connected. In the illustrative example in Fig. 7.2, v_1 is the centre node, while nodes $v_i, i = 2, \dots, 7$ are subject nodes, and node v_8 is the attacker node, forming edges e_{78}, e_{87} with node v_7 . Actually, there are a number of interesting ways to attack the social dominance of v_1 , and listed below are some of the most important/fundamental methods, which will be considered in this chapter. An example of each topology variation is provided in Fig. 7.2 to 7.6.

Topology 7.1 (Single Attack). Suppose that $n \geq 4$. Suppose further that $\bar{\mathcal{G}}[\bar{\mathcal{C}}]$ has star topology, with v_1 being the centre node, and with $n - 2$ subject nodes, $v_i, i = 2, \dots, n - 1$. A single attacker node v_n attaches to subject node v_{n-1} by forming edges $e_{n-1,n}$ and $e_{n,n-1}$. This forms the modified graph $\mathcal{G}[\mathcal{C}]$.

Topology 7.2 (Coordinated Double Attack). Suppose that $n \geq 5$. Suppose further that $\bar{\mathcal{G}}[\bar{\mathcal{C}}]$ has star topology, with v_1 being the centre node, and with $n - 3$ subject nodes, $v_i, i = 2, \dots, n - 2$. Two attacker nodes v_{n-1} and v_n attach to subject node v_{n-2} by forming the set of edges $\{e_{n-2,n-1}, e_{n-1,n-2}, e_{n-2,n}, e_{n,n-2}\}$. This forms $\mathcal{G}[\mathcal{C}]$.

Topology 7.3 (Uncoordinated Double Attack). Suppose that $n \geq 5$. Suppose further that $\bar{\mathcal{G}}[\bar{\mathcal{C}}]$ has star topology, with v_1 being the centre node, and with $n - 3$ subject nodes, $v_i, i = 2, \dots, n - 2$. One attacker node v_{n-1} attaches to subject node v_{n-3} with edges $e_{n-3,n-1}, e_{n-1,n-3}$. A second attacker node v_n attaches to subject node v_{n-2} with edges $e_{n-2,n}, e_{n,n-2}$. This forms $\mathcal{G}[\mathcal{C}]$.

Topology 7.4 (Two Dissenting Subjects). Suppose that $n \geq 4$. Suppose further that $\bar{\mathcal{G}}[\bar{\mathcal{C}}]$ has star topology, with v_1 being the centre node, $n - 1$ subject nodes $v_i, i = 2, \dots, n$. (There are no attacker nodes). Subject nodes v_{n-1} and v_n form edges $e_{n,n-1}, e_{n-1,n}$, forming $\mathcal{G}[\mathcal{C}]$.

The following topology variation is motivated by the concept of a *leadership group* where two leaders exist, and seek to maintain their collective social dominance.

Topology 7.5 (Leadership Group). Suppose that $\mathcal{G}_1[\mathbf{C}_1]$ and $\mathcal{G}_2[\mathbf{C}_2]$ respectively have $n \geq 3$ and $m \geq 3$ nodes, with node set $\mathcal{V}_1 = \{1, \dots, n\}$ and $\mathcal{V}_2 = \{n+1, \dots, n+m\}$ respectively. Both \mathcal{G}_1 and \mathcal{G}_2 have star topology; the centre nodes for \mathcal{G}_1 and \mathcal{G}_2 are v_1 and v_{n+1} respectively. Let $\mathcal{G}[\mathbf{C}]$ be the graph formed by merging \mathcal{G}_1 and \mathcal{G}_2 by insertion of the edges $e_{1,n+1}$ and $e_{n+1,1}$. Nodes v_1, v_{n+1} form a leadership group with subjects $v_2, \dots, v_n, v_{n+2}, \dots, v_{n+m}$.

Remark 7.3. Unlike Chapter 6, this chapter deals with the DeGroot–Friedkin model which has a constant \mathbf{C} for all issues $s \in \mathcal{S}$. The topology changes detailed above are assumed to have taken place before any discussion begins on issue $s = 0$. However, the exponential convergence conclusion obtained in Chapter 6 implies that for any given non-star network, there is a minimum convergence rate to the unique limiting trajectory $\mathbf{x}^*(s)$. Then, the results in this chapter are applicable if the topology changes at a rate that is sufficiently slow when compared to the minimum convergence rate of the network with dynamic topology.

In the next section, the above topological variations of the star graph are investigated. It is worth emphasising at this stage that, in Variations 7.1–7.3, it is useless for an attacker node v_n to attach to the centre node v_1 instead of a subject node; the topology remains a strongly connected star, and v_1 continues to be an autocrat. Note also that when new edges are introduced, it is assumed that each individual i adjusts weights c_{ij} to ensure that the new \mathbf{C} is row-stochastic. Take Topology Variation 7.5 as an example. Separately, the relative interaction matrix \mathbf{C}_1 (respectively \mathbf{C}_2) associated with \mathcal{G}_1 (respectively \mathcal{G}_2) are assumed to be row-stochastic. The relative interaction matrix \mathbf{C} associated with \mathcal{G} is also implicitly assumed to be row-stochastic with zero diagonal. That is, it is assumed that after the addition of edges $e_{1,n+1}$ and $e_{n+1,1}$, adjustments are made to the original weights $c_{1,j}$ and $c_{n+1,k}$ to ensure \mathbf{C} is row-stochastic.

Remark 7.4 (Ordering of Social Power). Theorem 7.1 states that \mathbf{x}^* is uniquely determined by γ , and Corollary 5.2 gives an upper bound on x_i^* given γ_i . On the other hand, Statement (ii) of Theorem 7.1 states that the ordering of x_i^* is consistent with the ordering of γ_i . This chapter is interested in the ordering of individual social power, as opposed to the precise values of social power, and this is reflected in Definition 7.3. By modifying the topology to affect the ordering of the γ_i , the work in this chapter can be considered as a problem in modifying topology to affect the eigenvector centralities of the nodes in the graph. Remark 5.1 provides more details on the importance of eigenvector centrality in network science problems.

7.3 Results and Their Social Interpretations

The analysis of each topology variation is now presented, with necessary and sufficient conditions established on the strength of the new relative interpersonal interactions (newly introduced edges) to ensure that individual 1 is no longer socially dominant. The social interpretations of these conditions are then discussed. To aid in presentation of the results and discussions, all proofs are deferred to the appendix of this chapter, Section 7.6.

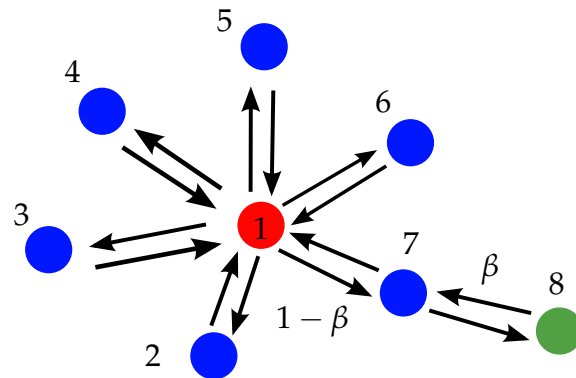


Figure 7.2: Topology Variation 7.1 (Single Attacker) with $n = 8$, attacker nodes are green.

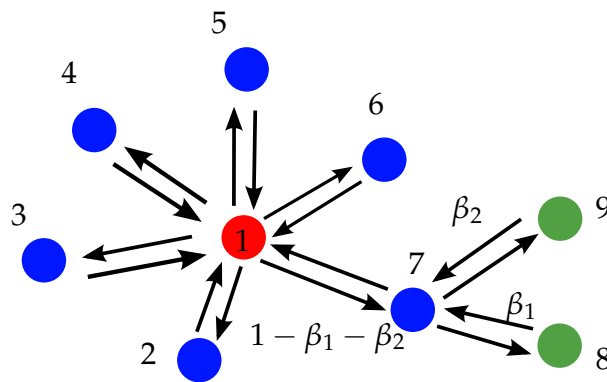


Figure 7.3: Topology Variation 7.2 (Coordinated Double Attacker) with $n = 9$, attacker nodes are green.

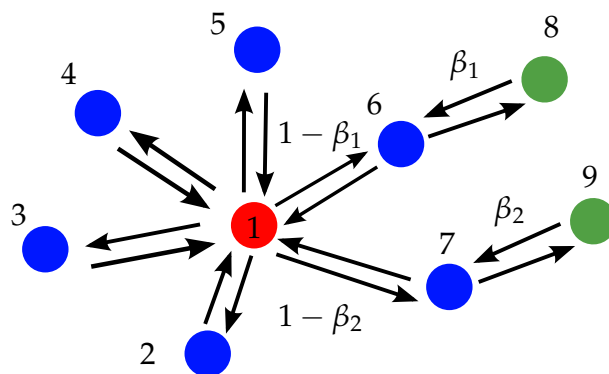


Figure 7.4: Topology Variation 7.3 (Uncoordinated Double Attacker) with $n = 9$.

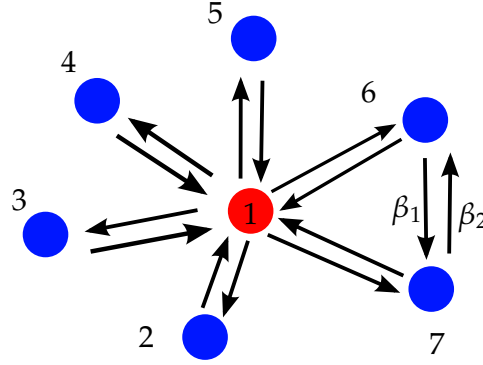


Figure 7.5: Topology Variation 7.4 (Two Dissenting Subjects) with $n = 7$.

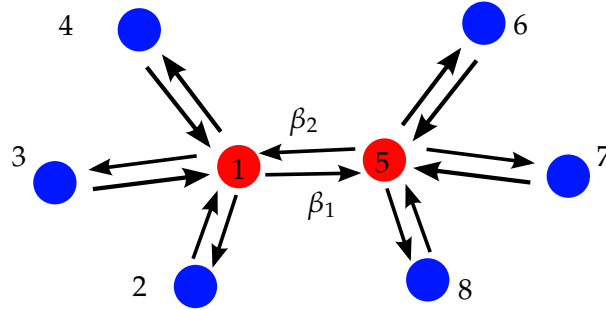


Figure 7.6: Topology Variation 7.5 (Leadership Group) with $n = m = 4$.

7.3.1 Topology Variation 7.1: Single Attack

First, consider Topology Variation 7.1. The relative interaction matrix $C(\beta)$ is

$$C(\beta) = \begin{bmatrix} 0 & c_{12} & c_{13} & \dots & c_{1,n-1} & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 - \beta & 0 & 0 & \dots & 0 & \beta \\ 0 & 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \quad (7.3)$$

where $\beta = c_{n-1,n} \in (0,1)$ is the influence² exerted by the attacker node v_n on subject node v_{n-1} . The following theorem details how the social power of each individual changes as a function of β . The exact form of C for other Topology Variations are provided in Appendix Section 7.6, as part of the proofs for the relevant results.

Theorem 7.2 (Single Attack). *Consider a social network with Topology Variation 7.1, with $\mathcal{G}[C]$ satisfying Assumption 7.1. Suppose that the individual social power vector $\mathbf{x}(s)$ evolves according to the DeGroot–Friedkin dynamics in Eq. (7.1), with initial conditions satisfying $0 \leq x_i(0) < 1, \forall i \in \mathcal{I}$ and for at least one $j \in \mathcal{I}$, $x_j(0) > 0$. Then, $\lim_{s \rightarrow \infty} \mathbf{x}(s) = \mathbf{x}^*$,*

²If $\beta = 0$ or $\beta = 1$ then \mathcal{G} is no longer strongly connected.

where $\mathbf{x}^* \in \text{int}(\Delta_n)$ is the unique fixed point of the map $F(\mathbf{x})$ in Eq. (7.2), and the following statements are true:

- (i) For all values of $\beta \in (0, 1)$, there holds $x_i^* < x_1^*$, for all $i \neq 1, n-1, n$ and $x_n^* < x_{n-1}^*$
- (ii) There holds 1) $x_1^* > x_i^*, \forall i \neq 1$ if and only if $\beta < 1 - c_{1,n-1} = \sum_{i=2}^{n-2} c_{1,i}$, or 2) $x_{n-1}^* > x_i^*, \forall i \neq n-1$ if and only if $\beta > 1 - c_{1,n-1}$. There holds $x_1^* = x_{n-1}^* > x_i^*, \forall i \neq 1, n-1$ if and only if $\beta_1 = 1 - c_{1,n-1}$
- (iii) There holds $x_n^* > x_1^*$ if and only if $\beta > 1/(1 + c_{1,n-1})$.

Corollary 7.1 (Generalised Placement of Single Attacking Node). Suppose that instead of attaching to subject node v_{n-1} , attacker node v_n can attach to any subject node $v_i, i \in \{2, \dots, n-1\}$ by forming edges $e_{n,i}, e_{i,n}$. The lower bound on $\beta = c_{n,i}$ required to have $x_i^* > x_1^*$ is minimised if v_n attaches to v_k where $k = \text{argmax}_{j \in \{2, \dots, n-1\}} c_{1,j}$.

Topology Variation 7.1 is in fact a special case of Topology Variation 7.2, obtained by setting $c_{n,n-2} = 0$ and by removing node v_n . The discussions on the social interpretation of the above results are deferred to after the statement of results on Variation 7.2, with discussions applicable to Variation 7.1 with minor adjustments.

7.3.2 Topology Variation 7.2: Coordinated Double Attack

Consider now Topology Variation 7.2. Let $\beta_1 = c_{n-2,n-1} \in (0, 1)$ and $\beta_2 = c_{n-2,n} \in (0, 1)$ be the two adjustable interpersonal weights accorded to the two attackers by subject v_{n-2} . Because \mathbf{C} is assumed to be row-stochastic, it is implied that $\beta_1 + \beta_2 + c_{n-2,1} = 1 \Rightarrow \beta_1 + \beta_2 < 1$ because $c_{n-2,1} > 0$. The following results hold.

Theorem 7.3 (Coordinated Double Attack). Consider a social network with Topology Variation 7.2, with $\mathcal{G}[\mathbf{C}]$ satisfying Assumption 7.1. Suppose that the individual social power vector $\mathbf{x}(s)$ evolves according to the DeGroot–Friedkin dynamics in Eq. (7.1), with initial conditions satisfying $0 \leq x_i(0) < 1, \forall i \in \mathcal{I}$ and for at least one $j \in \mathcal{I}$, $x_j(0) > 0$. Then, $\lim_{s \rightarrow \infty} \mathbf{x}(s) = \mathbf{x}^*$, where $\mathbf{x}^* \in \text{int}(\Delta_n)$ is the unique fixed point of the map $F(\mathbf{x})$ in Eq. (7.2), and the following statements are true:

- (i) For all $\beta_1, \beta_2 \in (0, 1)$, there holds $x_i^* < x_1^*$ for all $i \neq 1, n-2, n-1, n$, and $x_n^*, x_{n-1}^* < x_{n-2}^*$.
- (ii) There holds 1) $x_1^* > x_i^*, \forall i \neq 1$ if and only if $\beta_1 + \beta_2 < 1 - c_{1,n-2} = \sum_{i=2}^{n-3} c_{1,i}$, or 2) $x_{n-2}^* > x_i^*, \forall i \neq n-2$ if and only if $\beta_1 + \beta_2 > 1 - c_{1,n-2}$. There holds $x_1^* = x_{n-2}^*$ if and only if $\beta_1 + \beta_2 = 1 - c_{1,n-2}$.
- (iii) There holds $x_n^* > x_1^*$ (respectively $x_{n-1}^* > x_1^*$) if and only if $\beta_2 > (1 - \beta_1)/(1 + c_{1,n-2})$ (respectively $\beta_1 > (1 - \beta_2)/(1 + c_{1,n-2})$).
- (iv) There holds $x_{n-1}^* < x_n^*$ or $x_{n-1}^* > x_n^*$ if and only if $\beta_1 < \beta_2$ or $\beta_1 > \beta_2$ respectively. If $\beta_1 = \beta_2$, then $x_{n-1}^* = x_n^*$.

Corollary 7.2 (Generalised Placement of Coordinated Double Attack). *Suppose that instead of attaching to subject node v_{n-2} , attacker nodes v_{n-1}, v_n can attach to any subject node $v_i, i \in \{2, \dots, n-2\}$ by forming the set of edges $\{e_{n-1,i}, e_{i,n-1}, e_{n,i}, e_{i,n}\}$. The lower bound on $\beta_1 + \beta_2 = c_{n-1,i} + c_{n,i}$ required to have $x_i^* > x_1^*$ is minimised if v_{n-1} and v_n attach to v_k where $k = \operatorname{argmax}_{j \in \{2, \dots, n-2\}} c_{1,j}$.*

The above results can be interpreted in the following social context. One concludes from Statement (i) that subject node v_i , for $i \in \{2, \dots, n-3\}$, will never have greater social power at equilibrium than the centre individual v_1 , regardless of how β_1, β_2 change, i.e. $x_i^* < x_1^*$. Moreover, the attacker nodes v_{n-1}, v_n will never have greater social power than the subject node v_{n-2} to which it is attached.

Remark 7.2 states that c_{ij} can be considered the trust level accorded to individual j by individual i . The key result is Statement (ii), which indicates that the *combined trust* given to attackers v_{n-1} and v_n by subject node v_{n-2} must exceed the *combined trust* given to subjects v_2, \dots, v_{n-3} by centre node v_1 , in order for subject v_{n-2} to become socially dominant. It is most interesting to note that *it is only the sum* of the trust/interaction strength $\beta_1 + \beta_2$ that is relevant, and *there is no requirement on the individual magnitudes* of β_1, β_2 .

Corollary 7.2 delivers an intuitive, powerful, and socially relevant result. It states that the attackers v_{n-1}, v_n should seek to form an interpersonal relationship with the subject individual v_k that individual v_1 trusts the most. This minimises the required amount of trust subject v_k accords attacker v_n before centre v_1 loses social dominance.

Statement (iii) reveals that the attackers can also obtain social power greater than the centre individual v_1 if β_1 and β_2 exceed some lower bounding inequalities. Specifically, the inequality, which if satisfied ensures that attacker v_n has social power greater than centre v_1 , simply involves the variables β_1, β_2 and $c_{1,n-2}$. Moreover, there always exists a pair β_1, β_2 , satisfying $\beta_1 + \beta_2 < 1$ (and thus maintaining $\sum_{j=1}^n c_{n-2,j} = 1$), which ensures both attacker individuals v_{n-1}, v_n have social power greater than the centre v_1 . This is detailed in the proof in Section 7.6.2 and is desirable from the view point of the two attacker individuals, because it indicates that it is always possible to adjust the interpersonal interaction strengths of just β_1 and β_2 so that x_{n-1}^* and x_n^* are both greater than x_1^* .

Statement (iv) indicates that the ordering of the social powers x_{n-1}^* and x_n^* depends simply on the ratio β_1/β_2 . Thus, attacker v_{n-1} will have greater social power than attacker v_n if subject v_{n-2} trusts v_{n-1} more than v_n ($\beta_1 > \beta_2$) and vice versa.

7.3.3 Topology Variation 7.3: Uncoordinated Double Attack

Consider now Topology Variation 7.3, where two attackers attach onto two different subject nodes. Define $\beta_1 = c_{n-3,n-1} \in (0, 1)$ and $\beta_2 = c_{n-2,n} \in (0, 1)$, where $c_{n-3,n-1}$ and $c_{n-2,n}$ are the relative interpersonal weights given to attacker v_n by subject v_{n-2} and to attacker v_{n-1} by subject v_{n-3} , respectively. The following result is available.

Theorem 7.4 (Uncoordinated Double Attack). *Consider a social network with Topology Variation 7.3, with $\mathcal{G}[\mathcal{C}]$ satisfying Assumption 7.1. Suppose that the individual social power*

vector $\mathbf{x}(s)$ evolves according to the DeGroot–Friedkin dynamics in Eq. (7.1), with initial conditions satisfying $0 \leq x_i(0) < 1, \forall i \in \mathcal{I}$ and for at least one $j \in \mathcal{I}$, $x_j(0) > 0$. Then, $\lim_{s \rightarrow \infty} \mathbf{x}(s) = \mathbf{x}^*$, where $\mathbf{x}^* \in \text{int}(\Delta_n)$ is the unique fixed point of the map $\mathbf{F}(\mathbf{x})$ in Eq. (7.2), and the following statements are true:

- (i) For all $\beta_1, \beta_2 \in (0, 1)$, there holds $x_i^* < x_1^*$ for all $i \in \{2, \dots, n-4\}$, and $x_{n-1}^* < x_{n-3}^*$ and $x_n^* < x_{n-2}^*$.
- (ii) There holds $x_1^* > x_i^*$ for all $i \neq 1$ if and only if $\beta_1 < 1 - c_{1,n-3}$ and $\beta_2 < 1 - c_{1,n-2}$. If $\beta_1 > 1 - c_{1,n-3}$ (respectively $\beta_2 > 1 - c_{1,n-2}$), then $x_{n-3}^* > x_1^*$ (respectively $x_{n-2}^* > x_1^*$).
- (iii) For $i \in \{1, 2\}$, there holds $x_{n-2+i}^* > x_1^*$ if and only if $\beta_i > 1/(1 + c_{1,n-4+i})$.
- (iv) There holds $x_{n-3}^* > x_{n-2}^*$ if and only if $\frac{1-\beta_2}{1-\beta_1} > \frac{c_{1,n-2}}{c_{1,n-3}} \Leftrightarrow \frac{c_{1,n-3}}{c_{n-3,1}} > \frac{c_{1,n-2}}{c_{n-2,1}}$.

To reduce repetition, discussion is provided only to those results of Theorem 7.4 with social implications which differ significantly from those discussed previously regarding Theorem 7.3.

The most interesting conclusion drawn arises from comparing Theorem 7.4 to Theorem 7.3 (which concerns Topology Variation 7.2). With Topology Variation 7.2, it is required that the sum of the trust values $\beta_1 + \beta_2$ exceed a lower bound for the centre individual v_1 to lose social dominance, and there are no separate lower bounding inequalities for β_1 or β_2 . With Topology Variation 7.3, the centre individual v_1 loses social dominance if and only if either β_1 or β_2 exceed their respective lower bounding inequalities. Importantly, these two lower bounding inequalities are independent of each other. This shows that a **coordinated attack** on the social dominance of the centre node is more effective than an uncoordinated attack, an idea which is socially intuitive.

Moreover, from Statement (iii), it is clear that both attacker nodes have larger social power than the centre node if and only if $\beta_1 > 1/(1 + c_{1,n-3})$ and $\beta_2 > 1/(1 + c_{1,n-2})$, which implies that $\beta_1 + \beta_2 > 1/(1 + c_{1,n-3}) + 1/(1 + c_{1,n-2})$. From Statement (iii) of Theorem 7.3, with Topology Variation 7.2, both attacker nodes have larger social power than the centre node if and only if $\beta_2 > (1 - \beta_1)/(1 + c_{1,n-2})$ and $\beta_1 > (1 - \beta_2)/(1 + c_{1,n-2})$, which implies that $\beta_1 + \beta_2 > 2/(2 + c_{1,n-2})$. Since both $1 + c_{1,n-2}$ and $1 + c_{1,n-3}$ are smaller than $2 + c_{1,n-2}$, it follows that $1/(1 + c_{1,n-3}) + 1/(1 + c_{1,n-2}) > 2/(2 + c_{1,n-2})$, which implies that a **coordinated attack** on the social dominance of the centre node is also more efficient for the attackers, at least from the point of minimising $\beta_1 + \beta_2$ required for both attackers to have greater social power than the centre node.

7.3.4 Topology Variation 7.4: Two Dissenting Subjects

Topology Variation 7.4 is different from the ones studied above in the sense that there are no attacker nodes. Instead, one can consider this variation as one where two subjects form a relationship in *dissent* from the leader. Letting $\beta_1 = c_{n-1,n} \in (0, 1)$ and $\beta_2 = c_{n,n-1} \in (0, 1)$, analysis yields:

Theorem 7.5 (Two Dissenting Subjects). *Consider a social network with Topology Variation 7.4, with $\mathcal{G}[\mathbf{C}]$ satisfying Assumption 7.1. Suppose that the individual social power vector $\mathbf{x}(s)$ evolves according to the DeGroot–Friedkin dynamics in Eq. (7.1), with initial conditions satisfying $0 \leq x_i(0) < 1, \forall i \in \mathcal{I}$ and for at least one $j \in \mathcal{I}, x_j(0) > 0$. Then, $\lim_{s \rightarrow \infty} \mathbf{x}(s) = \mathbf{x}^*$, where $\mathbf{x}^* \in \text{int}(\Delta_n)$ is the unique fixed point of the map $F(\mathbf{x})$ in Eq. (7.2), and the following statements are true:*

- (i) *For all $\beta_1, \beta_2 \in (0, 1)$, there holds $x_i^* < x_1^*$ for all $i \neq 1, n-1, n$.*
- (ii) *There holds $x_n^* > x_1^*$ if and only if $\beta_1 > (1 - c_{1,n}) / (c_{1,n-1} + \beta_2)$ with $\beta_1 \in (0, 1)$. There exists such a $\beta_1 \in (0, 1)$ only if $\beta_2 > \sum_{i=2}^{n-2} c_{1,i}$.*
- (iii) *There holds $x_{n-1}^* > x_1^*$ if and only if $\beta_2 > (1 - c_{1,n-1}) / (c_{1,n} + \beta_1)$ with $\beta_2 \in (0, 1)$. There exists such a $\beta_2 \in (0, 1)$ only if $\beta_1 > \sum_{i=2}^{n-2} c_{1,i}$.*
- (iv) *There holds $x_n^* < x_{n-1}^*$ if and only if $\beta_2 > \beta_1 c_{1,n} + c_{1,n-1}(c_{1,n} - 1)$ or equivalently $\beta_1 < (\beta_2 + c_{n-1}(1 - c_{1,n}))$*

Note that the inequality in Statement (ii) can be rewritten as $\beta_2 > (1 - c_{1,n} - \beta_1 c_{1,n-1}) / \beta_1$ which, with $\beta_2 \in (0, 1)$, is satisfiable only if $\beta_1 > (1 - c_{1,n}) / (1 + c_{1,n-1})$. Similarly, the inequality in Statement (iii) is equivalent to $\beta_1 > (1 - c_{1,n-1} - \beta_2 c_{1,n}) / \beta_2$ which, with $\beta_1 \in (0, 1)$, is satisfiable only if $\beta_2 > (1 - c_{1,n-1}) / (1 + c_{1,n})$.

The interpretation for Statement (ii), which gives conditions on subject node v_n having greater social power than centre node v_1 , is now discussed. A similar conclusion can be drawn for Statement (iii), which concerns subject node v_{n-1} . Statements (ii) and (iii) are viewed as the key results for this specific topology variation. In order to make the centre node v_1 lose social dominance, the dissent subject nodes v_{n-1} and v_n must adopt a cooperative approach. From their definitions, one interprets β_1 as the trust given by v_{n-1} to v_n while β_2 is the trust given by v_n to v_{n-1} . A necessary condition for individual v_n to have social power greater than centre node v_1 is that $\beta_2 > \sum_{i=2}^{n-2} c_{1,i}$. This means that not only must v_{n-1} trust v_n sufficiently (as given by the inequality $\beta_1 > (1 - c_{1,n}) / (c_{1,n-1} + \beta_2)$), but individual v_n must reciprocate by ensuring that it trusts v_{n-1} sufficiently ($\beta_1 > \sum_{i=2}^{n-2} c_{1,i}$). Unless the two dissenting nodes build a cooperative and sufficiently strong bilateral relationship, centre node v_1 will remain socially dominant.

7.3.5 Topology Variation 7.5: Leadership Group

Last, Topology Variation 7.5 is considered. Here, the perspective shifts to the centre node of a star. Specifically, this topology variation considers the case where two star graphs are connected through the two centre nodes. The problem of interest is to obtain conditions on the relative interactions which guarantee that the two centre nodes continue to have greater social power than the subject nodes. With $\beta_1 = c_{1,n+1} \in (0, 1)$ and $\beta_2 = c_{n+1,1} \in (0, 1)$, the following result is obtained.

Theorem 7.6 (Leadership Group). *Consider a social network with Topology Variation 7.5, with $\mathcal{G}[\mathbf{C}]$ satisfying Assumption 7.1. Suppose that the individual social power vector*

$\mathbf{x}(s)$ evolves according to the DeGroot–Friedkin dynamics in Eq. (7.1), with initial conditions satisfying $0 \leq x_i(0) < 1, \forall i \in \mathcal{I}$ and for at least one $j \in \mathcal{I}$, $x_j(0) > 0$. Then, $\lim_{s \rightarrow \infty} \mathbf{x}(s) = \mathbf{x}^*$, where $\mathbf{x}^* \in \text{int}(\Delta_n)$ is the unique fixed point of the map $\mathbf{F}(\mathbf{x})$ in Eq. (7.2), and the following statements are true:

- (i) For all $\beta_1 \in (0, 1)$ and for all $\beta_2 \in (0, 1)$ there holds $x_i^* < x_1^*$ and $x_k^* < x_{n+1}^*$ for $i \in \{2, \dots, n\}$ and $k \in \{n+2, \dots, n+m\}$.
- (ii) There holds $x_1^* < x_{n+1}^*$ or $x_1^* > x_{n+1}^*$ if and only if $\beta_2 < \beta_1$ or $\beta_2 > \beta_1$ respectively. If $\beta_1 = \beta_2$, then $x_1^* = x_{n+1}^*$.
- (iii) For $k \in \{n+2, \dots, n+m\}$ there holds $x_k^* > x_1^*$ if and only if $c_{n+1,k}(\beta_1/\beta_2) > 1$. For $i \in \{2, \dots, n\}$, there holds $x_i^* > x_{n+1}^*$ if and only if $c_{1,i}(\beta_2/\beta_1) > 1$.

Statement (i) indicates that the centre individuals' social powers always remain higher than their respective subjects from the original separate star graphs. Statement (ii) shows that the ratio of β_1/β_2 determines whether centre node v_1 or centre node v_{n+1} is socially dominant. Statement (iii) reveals a surprising and interesting result on how leaders can *cooperatively* maintain collective social dominance over all subject nodes. Let $i \in \{2, \dots, n\}$ and $k \in \{n+2, \dots, n+m\}$, and approach from centre individual v_1 's point of view. While $x_i^* < x_1^*$ is guaranteed from Statement (i), in order to ensure that v_1 has greater social power than subject v_k (i.e. subjects of centre individual v_{n+1}), individual v_1 must ensure that $c_{n+1,k}(\beta_1/\beta_2) < 1$. This inequality always holds, regardless of the value of $c_{n+1,k} < 1$, if $\beta_1 = \beta_2$. That is, if the trust level v_1 accords to v_{n+1} is equal to the trust level v_{n+1} accords to v_1 , then regardless of the magnitude of this trust, v_1 has greater social power than *all* subject nodes *including the subjects of v_{n+1}* . Surprisingly, this holds even if $c_{n+1,k} \gg \beta_1, \beta_2$. Yet such a result is intuitive by considering β_1/β_2 as the ratio of the trust v_1 places on v_{n+1} (and thus indirectly the trust v_1 places on subject v_k via a directed path) versus the trust v_{n+1} places on v_1 (and thus indirectly the trust subject v_k places on v_1 via a directed path).

7.4 Simulations

In this section, some basic simulations are provided to illustrate selected results detailed in the previous section. First, the fundamental Topology Variation 7.1 is simulated. The simulation considers a network with $n = 8$ individuals (including the attacker individual). The top row of \mathbf{C} is $[0, 0.15, 0.15, 0.2, 0.05, 0.15, 0.3, 0]$. Figure 7.7 shows the social power at equilibrium x_i^* , as a function of $\beta = c_{78}$, for select individuals: centre individual v_1 , subject individuals v_2 and v_7 , and the attacker v_8 . As stated below Remark 7.3, for minor perturbations from $\beta = 0$, v_1 continues to be socially dominant. However, as β increases, x_1^* decreases while x_7^* and x_8^* increase. Centre v_1 loses social dominance when $\beta > 0.7$ (since $c_{17} = 0.3$) as stated in Theorem 7.2, Statement (ii). Moreover, $x_2^* < x_1^*$ and $x_8^* < x_7^*$ for all values of β as in Statement (i), and $x_8^* > x_1^*$ when β satisfies the inequality in Statement (iii).

Next, Topology Variation 7.4 is studied, in particular to show the need for cooperation between two dissenting individuals in order to ensure that the centre node

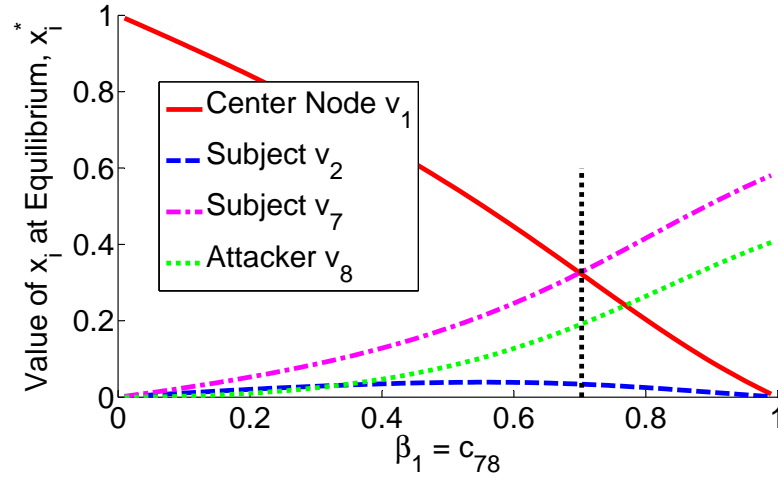


Figure 7.7: Simulation of Topology Variation 7.1 with $n = 8$, showing select individual $i = 1, 2, 7, 8$'s social powers at equilibrium, x_i^* as a function of the weight β , i.e. the relative interpersonal weight accorded to v_8 by v_1 . As β increases, x_1^* decreases. When $\beta > 1 - c_{17}$, v_1 is no longer the socially dominant individual in the network, as predicted by Theorem 7.2.

loses social dominance. The top row of C is $[0, 0.1, 0.1, 0.2, 0.05, 0.05, 0.2, 0.3]$. Figure 7.8 shows the social power at equilibrium x_i^* , for selected individuals $i = 1, 2, 7, 8$, as a function of $\beta_1 = c_{78}$ when $\beta_2 = 0.49$. Since $\beta_2 < \sum_{i=2}^{n-2} c_{1,i}$, one can verify that the necessary condition (the inequality on β_2) in Theorem 7.5, Statement (ii) is not satisfied. As a consequence, dissent subject v_8 never achieves social power greater than centre v_1 no matter how β_1 varies because there does not exist a $\beta_1 \in (0, 1)$ satisfying the required inequality $\beta_1 > (1 - c_{1,n}) / (c_{1,n-1} + \beta_2)$. Figure 7.9 shows the same simulation scenario but now with $\beta_2 = 0.55 > \sum_{i=2}^{n-2} c_{1,i}$, which satisfies the inequality condition on β_2 in Statement (ii). One can see, and in accordance with Theorem 7.5's Statement (ii), that $x_8^* > x_1^*$ when $\beta_1 > 0.93$. From individual v_8 's perspective, this emphasises the need for $c_{87} = \beta_2$ to be sufficiently large (i.e. v_8 has to trust v_7 enough) before change to $\beta_1 = c_{78}$ should be contemplated.

7.5 Conclusions

This chapter addressed a network modification problem for star topologies, which have the property that all social power accumulates with the centre individual. A number of strategies have been proposed for modifying a star graph via the insertion of new individuals and/or new interpersonal interactions. For each strategy, necessary and sufficient conditions are obtained on the strength of the new interpersonal interactions required for a change in social dominance. These conditions are interpreted from a social context, leading to several revealing conclusions. The Coordinated Double Attack strategy is more efficient and effective than the Uncoordinated

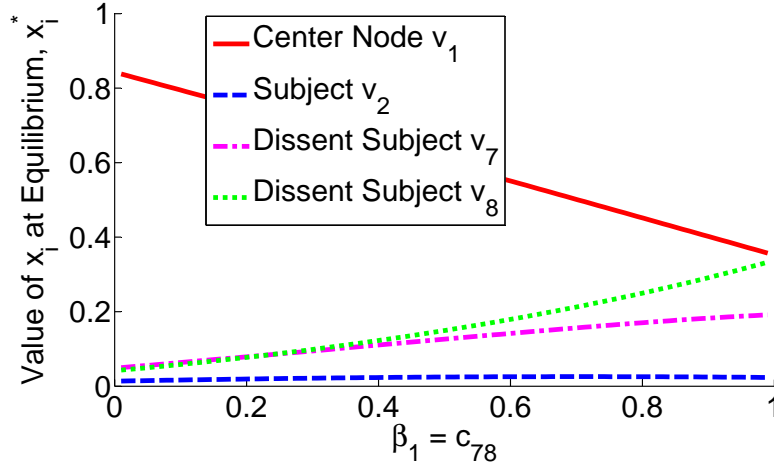


Figure 7.8: Simulation of Topology Variation 7.4 with $n = 8$, showing select individual $i = 1, 2, 7, 8$'s social powers at equilibrium, x_i^* as a function of the weight β_1 , i.e. the relative interpersonal weight accorded to v_8 by v_7 . In this instance, $\beta_2 < \sum_{i=2}^{n-2} c_{1,i}$, which does not satisfy the inequality on β_2 in Statement (ii), Theorem 7.5. No matter how β_1 varies, $x_1^* > x_i^*$ for all $i \neq 1$.

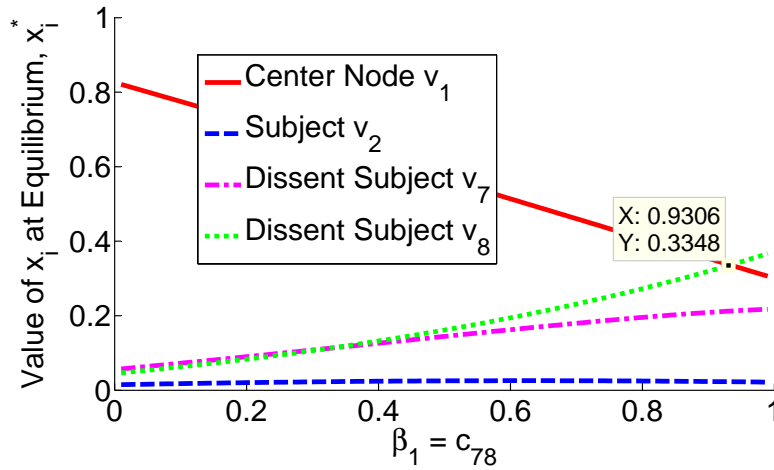


Figure 7.9: Simulation of Topology Variation 7.4 with $n = 8$, showing select individual $i = 1, 2, 7, 8$'s social powers at equilibrium, x_i^* as a function of the weight β_1 , i.e. the relative interpersonal weight accorded to v_8 by v_7 . In this instance, $\beta_2 > \sum_{i=2}^{n-2} c_{1,i}$, i.e. the inequality on β_2 in Statement (ii), Theorem 7.5 is satisfied. Thus, when β_1 is sufficiently large (see Statement (ii), Theorem 7.5), $x_1^* < x_8^*$ and centre v_1 is no longer the socially dominant individual.

Double Attack strategy. For the Dissenting Subjects strategy, cooperation among the two dissenting individuals is a necessary condition for the centre individual to lose social dominance. For the Leadership group, mutual trust between the two centre individuals (i.e. the ratio of the interpersonal weights the two centre individuals accord to each other must be close to 1) is a necessary and sufficient condition for the two centre individuals to maintain social dominance. All of the findings highlight a unifying concept: increasing (or decreasing) one's own social power is often only possible by modifying connections in other parts of the network.

7.6 Appendix: Proofs

The proofs extensively use Theorem 7.1, Statement (ii), which states that $x_i^* > x_j^* \Leftrightarrow \gamma_i > \gamma_j$ and $x_i^* = x_j^* \Leftrightarrow \gamma_i = \gamma_j$.

7.6.1 Proofs for Theorem 7.2 and Corollary 7.1

The expression $\gamma^\top = \gamma^\top C$, where C is given in Eq. (7.3), yields

$$\gamma_1 = \sum_{i=2}^{n-2} \gamma_i + (1 - \beta)\gamma_{n-1} \quad (7.4a)$$

$$\gamma_i = c_{1i}\gamma_1, \quad \forall i \neq 1, n-1, n \quad (7.4b)$$

$$\gamma_{n-1} = c_{1,n-1}\gamma_1 + \gamma_n \quad (7.4c)$$

$$\gamma_n = \beta\gamma_{n-1}. \quad (7.4d)$$

Statement (i) is obtained from Eq. (7.4b), where it is concluded that $\gamma_i < \gamma_1$ because $c_{1i} < 1$ for all $i \neq 1, n-1, n$, and from Eq. (7.4d), which allows one to conclude that $\gamma_n < \gamma_{n-1}$ for all $\beta \in (0, 1)$. For Statement (ii), substituting γ_n from Eq. (7.4d) into Eq. (7.4c) yields $\gamma_{n-1} = c_{1,n-1}\gamma_1 + \beta\gamma_{n-1}$. This is rearranged to obtain $\gamma_1 = \gamma_{n-1}(1 - \beta)/c_{1,n-1}$. Recalling that $0 < c_{1,n-1}$ and $0 < \beta < 1$, it follows that $\gamma_1 < \gamma_{n-1}$ if and only if $\beta > 1 - c_{1,n-1}$. Similarly, one can obtain that $\gamma_1 > \gamma_n$ if and only if $\beta < 1/(1 + c_{1,n-1})$, which proves Statement (iii). Corollary 7.1 is a generalisation of Statement (ii) obtained by observing that $\operatorname{argmin}_j(1 - c_{1,j}) = \operatorname{argmax}_j c_{1,j}$. \square

7.6.2 Proofs for Theorem 7.3 and Corollary 7.2

For Topology Variation 7.2, the relative interaction matrix C is given by

$$C(\beta) = \begin{bmatrix} 0 & c_{12} & c_{13} & \dots & c_{1,n-1} & 0 & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 - (\beta_1 + \beta_2) & 0 & 0 & \dots & 0 & \beta_1 & \beta_2 \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & 0 & \dots & 1 & 0 & 0 \end{bmatrix}. \quad (7.5)$$

From $\gamma^\top C = \gamma^\top$, it follows that

$$\gamma_1 = \sum_{i=2}^{n-3} \gamma_i + (1 - \beta_1 - \beta_2)\gamma_{n-2} \quad (7.6a)$$

$$\gamma_i = c_{1i}\gamma_1, \quad \forall i \neq 1, n-2, n-1, n \quad (7.6b)$$

$$\gamma_{n-2} = c_{1,n-2}\gamma_1 + \gamma_{n-1} + \gamma_n \quad (7.6c)$$

$$\gamma_{n-1} = \beta_1\gamma_{n-2} \quad (7.6d)$$

$$\gamma_n = \beta_2\gamma_{n-2}. \quad (7.6e)$$

Statement (i) is obtained from Eq. (7.6a) and Eq. (7.6d) and Eq. (7.6e), using the same arguments as in the proof for Theorem 7.2. Regarding Statement (ii), substitute Eq. (7.6d) and Eq. (7.6e) into Eq. (7.6c) and rearrange to obtain $\gamma_{n-2} = c_{1,n-2}\gamma_1 / (1 - \beta_1 - \beta_2)$. The statement is then straightforwardly obtained. For Statement (iii), in regards to γ_n , substitute $\gamma_{n-2} = c_{1,n-2}\gamma_1 / (1 - \beta_1 - \beta_2)$ into the right hand side of Eq. (7.6e) to obtain $\gamma_n = \beta_2 c_{1,n-2}\gamma_1 / (1 - \beta_1 - \beta_2)$. One can verify that $\beta_2 > (1 - \beta_1) / (1 + c_{1,n-2})$ implies $\beta_2 c_{1,n-2} / (1 - \beta_1 - \beta_2) > 1$, which in turn implies $\gamma_n > \gamma_1$. The inequality that ensures $\gamma_{n-1} > \gamma_1$ can be similarly found. Observe that $1 - \beta_1 < 1$, $1 - \beta_2 < 1$ and $1 < 1 + c_{1,n-2}$. There must also hold $\beta_1 + \beta_2 < 1$. This implies that for any value $c_{1,n-2}$, there always exist β_1, β_2 which ensures $\gamma_{n-1} > \gamma_1$ and $\gamma_n > \gamma_1$. From Eq. (7.6d) and Eq. (7.6e), one has that $\gamma_{n-1} / \gamma_n = \beta_1 / \beta_2$. Statement (iv) follows immediately. Corollary 7.2 is a generalisation of Statement (ii) by observing that $\arg\min_j (1 - c_{1,j}) = \arg\max_j c_{1,j}$. \square

7.6.3 Proof for Theorem 7.4

The equation $\gamma^\top C = \gamma^\top$, with C given by

$$C(\beta_1, \beta_2) = \begin{bmatrix} 0 & c_{1,2} & \cdots & c_{1,n-3} & c_{1,n-2} & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 1 - \beta_1 & 0 & \cdots & 0 & 0 & \beta_1 & 0 \\ 1 - \beta_2 & 0 & \cdots & 0 & 0 & 0 & \beta_2 \\ 0 & 0 & \cdots & 1 & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 & 0 \end{bmatrix}, \quad (7.7)$$

yields $\gamma_1 = (1 - \beta_1)\gamma_{n-3} + (1 - \beta_2)\gamma_{n-2} + \sum_{i=2}^{n-4} \gamma_i$, along with the equalities:

$$\gamma_i = c_i\gamma_1, \quad i \in \{2, \dots, n-4\} \quad (7.8a)$$

$$\gamma_{n-3} = c_{1,n-3}\gamma_1 + \gamma_{n-1} \quad (7.8b)$$

$$\gamma_{n-2} = c_{1,n-2}\gamma_1 + \gamma_n \quad (7.8c)$$

$$\gamma_{n-1} = \beta_1\gamma_{n-3} \quad (7.8d)$$

$$\gamma_n = \beta_2\gamma_{n-2} \quad (7.8e)$$

From (7.8a), since $c_{1,i} \in (0, 1)$ for all $i \in \{2, \dots, n-4\}$, it follows that $\gamma_i < \gamma_0$ for all $i \in \{2, \dots, n-4\}$. From (7.8d) and (7.8e), since $\beta_1, \beta_2 \in (0, 1)$, it follows that $\gamma_{n-1} < \gamma_{n-3}$ and $\gamma_n < \gamma_{n-2}$. Thus, Statement (i) is true.

From (7.8b) and (7.8d), one has that $\frac{\gamma_{n-3}}{\gamma_1} = \frac{c_{1,n-3}}{1-\beta_1}$, which implies that $\gamma_1 > \gamma_{n-3}$ if and only if $\beta_1 < 1 - c_{1,n-3}$. Similarly, from (7.8c) and (7.8e), one has that $\gamma_1 > \gamma_{n-2}$ if and only if $\beta_2 < 1 - c_{1,n-2}$. One then concludes that for $i \in \{1, 2\}$, if $\beta_i > 1 - c_{1,n-4+i}$, then $x_{n-4+i}^* > x_1^*$. Therefore, Statement (ii) is true.

From (7.8b) and (7.8d), one obtains that $\frac{\gamma_{n-1}}{\gamma_1} = \frac{\beta_1 c_{1,n-3}}{1-\beta_1}$. It follows that $\gamma_{n-1} > \gamma_1$ if and only if $\beta_1 > 1/(1 + c_{1,n-3})$. Similarly, from (7.8c) and (7.8e), one concludes that $\gamma_n > \gamma_1$ if and only if $\beta_2 > 1/(1 + c_{1,n-2})$. Thus, Statement (iii) is true.

Since $\frac{\gamma_{n-3}}{\gamma_1} = \frac{c_{1,n-3}}{1-\beta_1}$ and $\frac{\gamma_{n-2}}{\gamma_1} = \frac{c_{1,n-2}}{1-\beta_2}$, it follows that $\gamma_{n-3}(1 - \beta_1)/c_{1,n-3} = \gamma_{n-2}(1 - \beta_2)/c_{1,n-2}$, which implies that $\frac{\gamma_{n-3}}{\gamma_{n-2}} = \frac{c_{1,n-3}(1-\beta_2)}{c_{1,n-2}(1-\beta_1)}$. Then, $\gamma_{n-3} > \gamma_{n-2}$ if and only if $\frac{1-\beta_2}{1-\beta_1} > \frac{c_{1,n-2}}{c_{1,n-3}}$. Therefore, Statement (iv) is true. \square

7.6.4 Proof for Theorem 7.5

For Topology Variation 7.4, the relative interaction matrix \mathbf{C} is expressed as

$$\mathbf{C}(\beta_1, \beta_2) = \begin{bmatrix} 0 & c_{12} & c_{13} & \dots & c_{1,n-1} & c_{1,n} \\ 1 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 - \beta_1 & 0 & 0 & \dots & 0 & \beta_1 \\ 1 - \beta_2 & 0 & 0 & \dots & \beta_2 & 0 \end{bmatrix}, \quad (7.9)$$

where $\beta_1 = c_{n-1,n}$ and $\beta_2 = c_{n,n-1}$. The expression $\gamma^\top \mathbf{C} = \gamma^\top$ yields

$$\gamma_1 = \sum_{i=2}^{n-2} \gamma_i + (1 - \beta_1)\gamma_{n-1} + (1 - \beta_2)\gamma_n \quad (7.10a)$$

$$\gamma_i = c_{1i}\gamma_1, \quad \forall i \neq 1, n-1, n \quad (7.10b)$$

$$\gamma_{n-1} = c_{1,n-1}\gamma_1 + \beta_2\gamma_n \quad (7.10c)$$

$$\gamma_n = c_{1,n}\gamma_1 + \beta_1\gamma_{n-1}. \quad (7.10d)$$

Again, Statement (i) is obtained trivially from Eq. (7.10b). Substitute Eq. (7.10c) into Eq. (7.10d) and rearrange for γ_n to obtain

$$\gamma_n = \left(\frac{c_{1,n} + \beta_1 c_{1,n-1}}{1 - \beta_1 \beta_2} \right) \gamma_1, \quad (7.11)$$

and it follows that $\gamma_n > \gamma_1$ is implied by

$$c_{1,n} + \beta_1 c_{1,n-1} > 1 - \beta_1 \beta_2 \quad (7.12)$$

$$\beta_1 > \frac{1 - c_{1,n}}{c_{1,n-1} + \beta_2} \quad (7.13)$$

$$\beta_2 > \frac{1 - c_{1,n} - c_{1,n-1} \beta_1}{\beta_1}. \quad (7.14)$$

Consider Eq. (7.13). Observe that $(1 - c_{1,n}) / (c_{1,n-1} + \beta_2) \geq 1 \Leftrightarrow 1 - c_{1,n} - c_{1,n-1} \geq \beta_2 \Leftrightarrow \sum_{i=2}^{n-2} c_{1,i} \geq \beta_2$. Recalling that $\beta_1 \in (0, 1)$, one concludes that $\gamma_n > \gamma_1$ is possible only if $\beta_2 > \sum_{i=2}^{n-2} c_{1,i}$. Alternatively, one can consider Eq. (7.14) and similarly derive that $\gamma_n > \gamma_1$ if $\beta_2 > (1 - c_{1,n} - \beta_1 c_{1,n-1}) / \beta_1$ and $\beta_1 > (1 - c_{1,n}) / (1 + c_{1,n-1})$. The inequality conditions for ensuring $\gamma_{n-1} > \gamma_1$ are derived in a similar manner. \square

7.6.5 Proof for Theorem 7.6

The relative interaction matrix for Topology Variation 7.5 is given by

$$C(\beta_1, \beta_2) = \begin{bmatrix} 0 & c_{12} & c_{13} & \dots & \beta_1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \beta_2 & 0 & 0 & \dots & 0 & c_{n+1,n+2} & \dots & c_{n+1,n} \\ 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 & 0 & \dots & 0 \end{bmatrix} \quad (7.15)$$

And the expression $\gamma^\top C = \gamma^\top$ yields the following equalities

$$\gamma_1 = \sum_{1 < i \leq n} \gamma_i + \beta_2 \gamma_{n+1} \quad (7.16a)$$

$$\gamma_i = c_{1,i} \gamma_1, \quad \forall i \in \{2, \dots, n\} \quad (7.16b)$$

$$\gamma_{n+1} = \sum_{n+1 < i \leq n+m} \gamma_i + \beta_1 \gamma_1 \quad (7.16c)$$

$$\gamma_i = c_{n+1,i} \gamma_{n+1}, \quad \forall i \in \{n+2, \dots, n+m\} \quad (7.16d)$$

Statement (i) is obtained trivially from Eq. (7.16b) and Eq. (7.16d). In regards to Statement (ii), first substitute Eq. (7.16b) into Eq. (7.16a) to obtain $\gamma_1 = \beta_2 \gamma_{n+1} + \sum_{1 < i \leq n} c_{1,i} \gamma_1$ which is rearranged to yield $\gamma_1 (1 - \sum_{1 < i \leq n} c_{1,i}) = \beta_2 \gamma_{n+1}$, and which is equivalent to $\beta_1 \gamma_1 = \beta_2 \gamma_{n+1}$ because $1 - \sum_{1 < i \leq n} c_{1,i} = \beta_1$. Statement (iii) is obtained by substituting $\gamma_1 = \beta_2 \gamma_{n+1} / \beta_1$ into Eq. (7.16d). \square

Nonlinear Mapping Convergence and Application to Social Power Analysis

Chapter Summary

This chapter considers the original DeGroot–Friedkin model. Recently, a nonlinear mapping convergence result was developed using Lefschetz fixed point theory; a map can be shown to have a unique fixed point that is locally exponentially stable if the Jacobian evaluated at any fixed point satisfies a certain property. This result is used to analyse the DeGroot–Friedkin model, allowing an exponential convergence result to be established using a tool different to the nonlinear contraction analysis employed in Chapters 5 and 6.

It must be noted that both the general result and the application to the DeGroot–Friedkin model will appear in the same conference paper at the 2018 European Control Conference (see page vii, paper [10]). The general result was developed by B. D. O. Anderson, and will be presented in Section 8.2. The application to the DeGroot–Friedkin model was the work of the author, and will thus be presented separately.

8.1 Introduction

Recursive equations of the form

$$\mathbf{x}(k+1) = \mathbf{F}(\mathbf{x}(k)), \quad (8.1)$$

are fundamental to control and signal processing. Very often \mathbf{F} is linear or affine, but this is not always so and as a result, established linear system theory techniques cannot be applied. However, and as is the case in this chapter, \mathbf{F} is often suitably smooth. Usually also, $\mathbf{x}(k)$ resides in a Euclidean space of known dimension, though this is not always the case, and indeed will not always be the case in the problem considered in this chapter.

In many situations, it is possible to examine local behaviour of the nonlinear map

F in Eq. (8.1) around an equilibrium point, through a linearisation process. If \bar{x} is an equilibrium point, i.e. a fixed point of the mapping F satisfying $\bar{x} = F(\bar{x})$, then the Jacobian $J(\bar{x}) = \frac{\partial F}{\partial x}|_{\bar{x}}$ can often provide guidance as to behaviour in the vicinity of \bar{x} . If $\|x(k) - \bar{x}\|$ is small, for some norm $\|\cdot\|$, then approximately

$$x(k+1) - \bar{x} = J(\bar{x})[x(k) - \bar{x}] \quad (8.2)$$

If the eigenvalues of $J(\bar{x})$ do not lie on the unit circle, then the asymptotic stability or instability of the linear equation Eq. (8.2) implies the same property for the nonlinear equation Eq. (8.1), albeit locally.

The DeGroot–Friedkin model, introduced in [Jia et al., 2015] and studied in Chapters 5–7 thus far, is what amounts to a particular version of Eq. (8.1). [Jia et al., 2015] established by two different sets of rather specialised calculations, tailored to the specific algebraic form of F , that under normal circumstances, the system Eq. (8.1) (i) possesses a unique fixed point for the map F , and (ii) the equilibrium is globally asymptotically stable. Chapter 5 used nonlinear contraction theory to obtain an exponential convergence result, again exploiting the specific algebraic form of F . In that instance, the specific algebraic form of F was used to find the expression of $J(x)$ over the entire state-space, and from this expression of the Jacobian, a differential coordinate transform was introduced to establish a contractive property.

It is natural to speculate whether the conclusion that there is a single attractive equilibrium is indeed intrinsic to the algebraic form of F , or whether rather, it is a consequence of some more general property, and consequently also one that will follow for a whole class of F of which that in the DeGroot–Friedkin model is just a special case. It turns out that this conclusion is indeed a consequence of a more general property. The recent paper [Anderson and Ye, 2018] showed, using Lefschetz fixed point theory, that if a system of the form Eq. (8.1) acting on a positive invariant set (with the set satisfying some further topological conditions) has at least one fixed point and the Jacobian at every fixed point has eigenvalues inside the unit circle, then the map F has a unique fixed point and Eq. (8.1) is locally exponentially stable about that point. This chapter shows that the DeGroot–Friedkin model has a map F which is in the larger class of maps for which the result in [Anderson and Ye, 2018] holds.

By way of brief background, Lefschetz fixed point theory (of which more details are summarised subsequently) is a tool for relating the local behaviour of maps to some global properties, taking into account the underlying topological space in which the maps act. The local properties are associated with the linearised equations Eq. (8.2), potentially studied at multiple equilibrium points (and with in general a different $J(\bar{x})$ associated with each equilibrium point). Such local properties were flagged in Chapters 5 and 6 as of central concern in, respectively, a time-invariant and a time-varying version of the DeGroot–Friedkin model.

To sum up the contribution of this chapter, an exponential convergence result is established for the DeGroot–Friedkin model using a recently developed nonlinear mapping convergence result. The analysis method avoids the detailed calculations required in [Jia et al., 2015] and Chapter 5 to establish uniqueness and attractiveness

of the equilibrium, by only requiring evaluation of the Jacobian at *each fixed point* of the map (assuming at least one fixed point exists); uniqueness and local exponential stability are guaranteed if the Jacobian at each fixed point satisfies a certain property.

8.1.1 Chapter Organization

In the remaining part of this chapter, Section 8.2 reviews a recently developed nonlinear mapping result, and Section 8.3 shows how this result may be applied to the DeGroot–Friedkin model. Conclusions are drawn in Section 8.4.

8.2 A General Nonlinear Mapping Convergence Result

8.2.1 Background on Lefschetz Fixed Point Theory

Lefschetz fixed-point theory applies to smooth maps $F : X \rightarrow X$ where X is a compact oriented manifold [Guillemin and Pollack, 2010; Hirsch, 2012] or a compact triangulable space [Armstrong, 2013]¹. Thus $X = \mathbb{R}^n$ is excluded, but if X is a compact subset of \mathbb{R}^n such as a simplex, then it is allowed. This also means that if a map $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is known to have no fixed points for large values of its argument, the theory can often be applied by considering the restriction of F to a compact subset of \mathbb{R}^n , such as a ball of large enough radius for which F in Eq. (8.1) needs to be positively invariant.

Lefschetz fixed-point theory involves derivatives. Any smooth map has the property that at any point $x \in X$, there is a linear derivative mapping, call it dF_x , and if X looks locally like \mathbb{R}^m , then the derivative map can be represented by the $m \times m$ Jacobian matrix in the local coordinate basis.

Interest is centred for the purposes of this chapter on those maps which have a finite number of fixed points (including possibly zero) in X , though of course, maps with an infinite set of fixed points exist, for example $F(x) = x$, i.e. the identity map. A fixed point x is called a *Lefschetz fixed point* of F if the eigenvalues of dF_x are unequal to 1. A fixed point being a Lefschetz fixed point is sufficient but not necessary to ensure that x is an isolated fixed point of F , i.e. there is an open neighbourhood around x in which no other fixed point occurs. Because X is compact, and if it is known that all fixed points of F are isolated, say because they are all Lefschetz fixed points, it easily follows that the number of fixed points is necessarily finite. For completeness, an argument by contradiction is recorded, which seems standard. If there were an infinite number of fixed points, $x_i, i = 1, 2, \dots$, compactness of X implies there is a convergent subsequence x_{i_1}, x_{i_2}, \dots , with limit point \bar{x} , and again by compactness $\bar{x} \in X$. Now F is continuous so $F(x_{ij}) \rightarrow F(\bar{x})$ since $x_{ij} \rightarrow \bar{x}$ as $j \rightarrow \infty$. Then $x_{ij} - F(x_{ij}) \rightarrow \bar{x} - F(\bar{x})$ as $j \rightarrow \infty$. Since $x_{ij} - F(x_{ij}) = 0 \forall j$, it is evident

¹The notion of orientation of a manifold is described in the references; roughly, a manifold is oriented if one can attach an infinitesimal set of coordinate axes to an arbitrary point on the manifold, and then move the point with the axes attached knowing that one can never move to reverse the orientation. A Möbius strip is *not* an oriented manifold.

that \bar{x} is a fixed point of F . However, being a limit point it is not isolated, hence the contradiction.

The Lefschetz property holding at a particular fixed point x also implies that at the point x , the (linear) mapping $I_m - dF_x$ is an isomorphism of the tangent space $T_x(X)$ at x . If it preserves orientation, then its determinant is positive, while if it reverses orientation, its determinant is negative. The *local Lefschetz number* of F at a fixed point x , written $L_x(F)$, is defined as $+1$ or -1 according as the determinant of $I_m - dF_x$ is positive or negative.²

The map F is termed a *Lefschetz map* if and only if all its fixed points are Lefschetz fixed points (and there are then, as noted above, a finite number of fixed points). The *Lefschetz number* of F , written $L(F)$, is defined as

$$L(F) = \sum_{F(x)=x} L_x(F) \quad (8.3)$$

There is an alternative definition of the Lefschetz number not provided here which can be shown to be equivalent to that appearing here, based on topological considerations, and provided in [Hirsch, 2012; Guillemin and Pollack, 2010]. It is *not* restricted to maps with a finite number of fixed points. Moreover, using this alternative definition, one sees that $L(F)$ is a *homotopy invariant*³, and this particular property does not require limitation to those maps with a finite number of fixed points. Further, the alternative approach yields a connection between the Lefschetz number of the identity map (which has an infinite number of fixed points) and another topological invariant of the underlying space X , viz the Euler characteristic⁴, [Hirsch, 2012; Guillemin and Pollack, 2010; Matsumoto, 2002].

The key result (see e.g. [Hirsch, 2012] for the case of a compact oriented manifold and [Armstrong, 2013] for the case of a compact triangulable space) is as follows:

Theorem 8.1. *The Lefschetz number of the identity map $\mathcal{I}_d : X \rightarrow X$ where X is a compact oriented manifold or a compact triangulable space is $\chi(X)$, the Euler characteristic of X .*

A key consequence of this theorem is that if a map F is homotopically equivalent to \mathcal{I}_d , i.e. if there exists a smooth map $H : X \times I \rightarrow X$ such that $H(x, 0) = F(x)$ and $H(x, 1) = \mathcal{I}_d(x) = x$ then

$$L(F) = \chi(X) \quad (8.4)$$

Hence one has the following theorem:

²Reference [Guillemin and Pollack, 2010] uses $dF_x - I_m$ rather than $I_m - dF_x$, which is used by [Hirsch, 2012]. The latter form is what is required.

³Smooth maps $F : X \rightarrow X$ and $G : X \rightarrow X$ are said to be homotopic if there exists a smooth map $H : X \times I \rightarrow X \times I$ with $H(x, 0) = F(x)$, and $H(x, 1) = G(x)$. Saying $L(F)$ is a homotopy invariant means $L(F) = L(G)$ for any G which is homotopic to F .

⁴The Euler characteristic is an integer number associated with a topological space, including a space that in some sense is a limit of a sequence of multidimensional polyhedra, e.g. a sphere, and a key property is that distortion or bending of the space leaves the number invariant. Euler characteristics are known for a great many topological spaces.

Theorem 8.2 (Specialisation of Lefschetz–Hopf Theorem). *Let X be a compact oriented manifold or a compact triangulable space, and suppose $F : X \rightarrow X$ is a Lefschetz map, i.e. there are a finite number of fixed points at each of which $I_m - dF_x$ is an isomorphism, and is homotopically equivalent to the identity map. Then there holds*

$$L(F) = \sum_{F(x)=x} L_x(F) = \chi(X) \quad (8.5)$$

where $L_x(F)$ is $+1$ or -1 according as $\det(I_m - dF_x)$ has positive or negative sign, and $\chi(X)$ is the Euler characteristic of X .

8.2.2 A General Convergence Result

The general result established in [Anderson and Ye, 2018] will now be presented. It is assumed throughout this subsection that a mapping F has at least one fixed point (this perhaps being established by a standard fixed point result, e.g. Brouwer Fixed Point Theorem, [Guillemin and Pollack, 2010; Hirsch, 2012]). The result will establish that certain properties of the mapping F and the associated space X guarantee that F has a unique fixed point and the system Eq. (8.1) is locally exponentially stable about the unique fixed point. The main result, proved using the Lefschetz theory, is as follows.

Theorem 8.3. *Consider a smooth map $F : X \rightarrow X$ where X is a compact, oriented and convex manifold or a convex triangulable space of arbitrary dimension. Suppose that the eigenvalues of dF_x have magnitude less than 1 for all fixed points of F . Then F has a unique fixed point, and in a local neighborhood about the fixed point, Eq. (8.1) converges to the fixed point exponentially fast.*

Proof. Observe first that the compactness and convexity properties of X guarantee it is homotopy equivalent to the unit m -dimensional disk D^m and accordingly then homotopy equivalent to a single point. This means that $\chi(X) = 1$, see e.g. [Matsumoto, 2002, pp. 140].

Next, observe that, because X is convex, $H = t\mathcal{I}_d + (1-t)F$ which maps x to $tx + (1-t)F(x)$, is a mapping from X to X for every $t \in [0, 1]$ and the smoothness properties of H (which come from the smoothness of F and the specific dependence on t) then guarantee that F and \mathcal{I}_d are homotopically equivalent. By Theorem 8.2, there holds

$$L(F) = 1 \quad (8.6)$$

Now for any real matrix $A \in \mathbb{R}^{m \times m}$ for which the eigenvalues are less than one in magnitude, it is easily seen that the matrix $I_m - A$ has eigenvalues all with positive real part, from which it follows that the determinant of $I_m - A$ is positive, since the determinant is equal to the product of the eigenvalues. Hence for any fixed point x of F , one observes by identifying A with dF_x that there necessarily holds $L_x(F) = 1$. By Eq. (8.3) and Eq. (8.6), it follows that $1 = \sum_{F(x)=x} 1$, or that there is precisely one fixed point.

Convergence of Eq. (8.1) to the unique fixed point from any initial value in its region of attraction is necessarily exponentially fast. In a neighbourhood \mathcal{D} around the unique fixed point, the eigenvalue property of dF_x guarantees exponential convergence. The region of attraction for the fixed point is in most instances larger than \mathcal{D} , and one denotes as $\mathcal{U} \subset X$ an arbitrary compact space within the region of attraction, and containing \mathcal{D} . For any initial $x \in \mathcal{U}$, the sequence $x, F(x), F(F(x)), \dots$ converges to the neighbourhood \mathcal{D} in a finite number of steps, and because the set \mathcal{U} is compact, there is a number of steps, $\bar{N} < \infty$ say, such that from all initial conditions in \mathcal{U} , the neighbourhood is reached in no more than \bar{N} steps. The finiteness of \bar{N} then implies that exponentially fast convergence occurs $\forall x(0) \in \mathcal{U}$. \square

Remark 8.1. *The proof of the theorem using Lefschetz ideas will clearly generalise in the following way. Suppose that F is homotopic to the identity and X is not homotopic to the unit ball, while all fixed points are Lefschetz with the property that $I_m - dF_x$ has positive determinant. Then the number of fixed points will be $\chi(X)$, and Eq. (8.1) will have $\chi(X)$ locally exponentially stable equilibria. If for example F mapped S^2 to S^2 (i.e. the unit sphere embedded in three-dimensional Euclidean space) and never mapped a point to its antipodal point, i.e. there was no x for which $F(x) = -x$, it will be homotopic to the identity map and then there will be two fixed points, since $\chi(S^2) = 2$. To construct the homotopy, observe that, because of the exclusion that F can map any point to an antipodal point, there is a well-defined homotopy provided by*

$$H(x, t) = \frac{(1-t)x + tF(x)}{\|(1-t)x + tF(x)\|_2}$$

8.3 Application to the DeGroot-Friedkin Model

The above results are now applied to the DeGroot–Friedkin model. For simplicity, consider networks with $n \geq 3$ individuals. The model itself was introduced in Chapter 5 and its properties (including convergence) studied in Chapters 5 through 7. Thus, no further details about the model itself will be presented, though the map F in question is presented again, for convenience, and its properties discussed below.

$$F(x) = \begin{cases} \mathbf{e}_i & \text{if } x = \mathbf{e}_i \text{ for any } i \\ \alpha(x) \begin{bmatrix} \frac{\gamma_1}{1-x_1} \\ \vdots \\ \frac{\gamma_n}{1-x_n} \end{bmatrix} & \text{otherwise} \end{cases} \quad (8.7)$$

with $\alpha(x) = 1 / \sum_{i=1}^n \frac{\gamma_i}{1-x_i}$ where the vector $\gamma = [\gamma_1, \gamma_2, \dots, \gamma_n]^\top$ is constant, has strictly positive entries $\gamma_i \leq 1/2$ and satisfies $\gamma^\top \mathbf{1}_n = 1$. It can be verified that $F : \Delta_n \rightarrow \Delta_n$ where $\Delta_n = \{x_i : \sum_i x_i = 1, x_i \geq 0\}$ is the n -dimensional unit simplex (see Section 2.1). Thus, Δ_n satisfies all the requirements on compactness, orientability, and convexity. Moreover, F is smooth everywhere on Δ_n , including at the corners

$x_i = 1$, even given the $1/(1 - x_i)$ term in the i^{th} entry of F ; it was shown in the proof of Corollary 5.1 that F is of class \mathcal{C}^∞ in Δ_n . It was also proved in Chapter 5 (and also in [Jia et al., 2015]) that there are no fixed points of F for which there exists a j such that $x_j = 0$, other than the corners of the simplex at $x = \mathbf{e}_k, k = 1, \dots, n$. These corners were proved to be unstable fixed points unless $\gamma_i = 1/2$ (there can only be one entry of γ equal to $1/2$ because $n \geq 3$ and $\gamma_k > 0$ for all k), see Corollary 5.1. The case of $\gamma_i = 1/2$ occurs only for star graphs, which are not considered in this chapter.

One should also take note of the important point that the above definition Eq. (8.7) of F can be regarded as defining a map $\mathbb{R}^n \rightarrow \mathbb{R}^n$, or as defining a map on an $(n - 1)$ -dimensional triangulable space $\Delta_n \rightarrow \Delta_n$, with the n -dimensional vector $x = [x_1, x_2, \dots, x_n]^\top$ providing a convenient parametrization of the space given imposition of the constraints $\sum_{i=1}^n x_i = 1, x_i \geq 0$.

8.3.1 Existing Results

Before presenting the application of Theorem 8.3 to the DeGroot–Friedkin model, a quick recapitulation of the existing results on the DeGroot–Friedkin model (including those in the previous chapters of this thesis) are provided here to better contextualise the result in this chapter.

In the paper [Jia et al., 2015], which first proposed the DeGroot–Friedkin model, the following analysis was provided. First, because F is continuous and Δ_n is convex and compact, Brouwer’s Fixed Point Theorem is used to conclude there exists at least one interior fixed point. Next, [Jia et al., 2015] used a series of inequality calculations, exploiting the algebraic form of F , to show that the fixed point \bar{x} is unique, and importantly, that \bar{x} is in the interior of Δ_n . Following this, the authors showed using a separate set of complicated calculations that the trajectories of $x(k)$ had specific monotone properties, again by exploiting the algebraic form of F . Last, a Lyapunov function is proposed and the properties of the trajectories of $x(k)$ are used to show the Lyapunov function is nonincreasing; LaSalle’s Invariance Principle is used to conclude asymptotic convergence to the unique interior fixed point \bar{x} for all initial conditions $x(0)$ that are not a corner of the simplex Δ_n .

Chapter 5 takes a different approach, and looks at the Jacobian of F both as a map $\mathbb{R}^n \rightarrow \mathbb{R}^n$ and its restriction (after choice of an appropriate coordinate basis for Δ_n) as a map $\Delta_n \rightarrow \Delta_n$. (Note that in any fixed coordinate basis, the second Jacobian is of dimension $(n - 1) \times (n - 1)$, with the two Jacobians necessarily related, as described further below. It is this second Jacobian which represents the mapping dF_x defined in earlier sections.) However, rather than employing Theorem 8.3, Chapter 5 uses nonlinear contraction analysis. A differential transformation is introduced, with the transformation exploiting the algebraic form of F (and specifically the form of the relevant Jacobian). More importantly, that chapter studies the Jacobian over all values of x in the state-space (in that instance, x was in the interior of Δ_n). As will be apparent, application of Theorem 8.3 requires evaluation of the mapping dF_x only at the fixed points of F (which are not assumed to be unique). Chapter 5 concludes by

showing that the 1-norm of the transformed Jacobian is less than one via a nontrivial calculation, and thus exponential convergence to a unique fixed point is ensured, for all $\mathbf{x}(0)$ not at the corners of the simplex Δ_n .

8.3.2 Proof of a Unique Fixed Point Which Is Locally Exponentially Stable

Before the application of Theorem 8.3 to the DeGroot–Friedkin model is shown, establishing there is a single fixed point, and further that it is locally exponentially stable, the Jacobian of $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$, and the related Jacobian of $F : \Delta_n \rightarrow \Delta_n$ in a coordinate basis to be defined, are computed, and some properties of the two Jacobians are established. For convenience, and when there is no risk of confusion, the argument k is dropped from $\mathbf{x}(k)$ and $\mathbf{x}(k)$ from $\alpha(\mathbf{x}(k))$. Recall from Chapter 5 that

$$\frac{\partial F_i}{\partial x_i} = \frac{\gamma_i \alpha}{(1 - x_i)^2} - \frac{\gamma_i^2 \alpha^2}{(1 - x_i)^3} = F_i \frac{1 - F_i}{1 - x_i}. \quad (8.8)$$

Similarly, one obtains, for $j \neq i$,

$$\frac{\partial F_i}{\partial x_j} = -\frac{\gamma_i \gamma_j \alpha^2}{(1 - x_i)(1 - x_j)^2} = -\frac{F_i F_j}{1 - x_j}. \quad (8.9)$$

The above is in fact the Jacobian of $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Regarding the unstable equilibria at the corners of the simplex, the proof of Corollary 5.1 showed that the Jacobian $\frac{\partial F}{\partial \mathbf{x}}$ associated with $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ at the point $\mathbf{x} = \mathbf{e}_i$ has a single eigenvalue at $(1 - \gamma_i)/\gamma_i$ and all other eigenvalues are 0. Since $\gamma_i < 1/2$, then $(1 - \gamma_i)/\gamma_i > 1$ and the fixed point \mathbf{e}_i is unstable. The associated eigenvector is \mathbf{e}_i . This eigenvector has a nonzero projection onto Δ_n so that the instability is also an instability of the fixed point of $F : \Delta_n \rightarrow \Delta_n$. No matter what $(n - 1)$ -vector coordinatisation is used for Δ_n , the representation of $dF_{\mathbf{x}}$ will be an $(n - 1) \times (n - 1)$ matrix with an eigenvalue greater than 1.

Recall that there are no fixed points $\bar{\mathbf{x}}$ for which there is a zero entry in $\bar{\mathbf{x}}$, except the fixed points at the corners of the simplex. (As noted above, Brouwer's Fixed Point Theorem can be used to easily conclude that there is at least one such fixed point $\bar{\mathbf{x}}$.) Since \mathbf{e}_i for all $i = 1, \dots, n$ are unstable equilibria, they can be excluded by defining an entity⁵, distinct from Δ_n , as $\hat{\Delta}_n = \{\mathbf{x} : \sum_i^n x_i = 1, 0 < \delta \leq x_i \leq 1 - \delta\}$, where $\delta > 0$ is sufficiently small to ensure that any fixed point of F in Δ_n , save the unstable \mathbf{e}_i , is contained in $\hat{\Delta}_n$. This ensures that $\hat{\Delta}_n$ is a compact, convex, and oriented manifold, and it is easily verified to be positively invariant for Eq. (8.7) (see e.g. Chapter 5, Property 5.2). This allows the use of results developed in Section 8.2.2. In other words, one may now study the map in Eq. (8.7) as $F : \hat{\Delta}_n \rightarrow \hat{\Delta}_n$. Now as already noted the above computed $n \times n$ Jacobian $\frac{\partial F}{\partial \mathbf{x}}$, with elements given in

⁵This definition of $\hat{\Delta}_n$ in this chapter is slightly different to $\tilde{\Delta}_n$ defined in Section 2.1 and used in Chapters 5 through 7. The purpose of defining $\hat{\Delta}_n$ in this chapter in the manner described is to define a positive invariant set for the system Eq. (8.1) that is compact, convex and orientable, so Theorem 8.3 may be used.

Eq. (8.8) and Eq. (8.9), is in fact not what is required to apply⁶ Theorem 8.3. This is because $\frac{\partial F}{\partial x}$ is the Jacobian computed in the coordinates of the Euclidean space in which $\hat{\Delta}_n$ is embedded. What is instead required is the Jacobian *on the manifold* $\hat{\Delta}_n$, which will now be obtained. A new coordinate basis is introduced, $\mathbf{y} \in \mathbb{R}^{n-1}$ where $y_1 = x_1, y_2 = x_2, \dots, y_{n-1} = x_{n-1}$, and thus *on the manifold* $\hat{\Delta}_n$ there holds $x_n = 1 - \sum_{k=1}^{n-1} y_k$. On the manifold, and in the new coordinates, define \mathbf{G} as the map with $G_1(\mathbf{y}) = F_1(\mathbf{x}), \dots, G_{n-1}(\mathbf{y}) = F_{n-1}(\mathbf{x})$, which means that $F_n = 1 - \sum_{k=1}^{n-1} G_k$. The Jacobian on the manifold of $\hat{\Delta}_n$ is in fact $d\mathbf{G}_y$, which is now computed. For any $G_i(y_1, \dots, y_{n-1}) = F_i(y_1, \dots, y_{n-1}, 1 - \sum_{k=1}^{n-1} y_k)$, one obtains by the Chain rule that:

$$\frac{\partial G_i}{\partial y_j} = \sum_{k=1}^n \frac{\partial F_i}{\partial x_k} \frac{\partial x_k}{\partial y_j} = \frac{\partial F_i}{\partial x_j} \frac{\partial x_j}{\partial y_j} + \frac{\partial F_i}{\partial x_n} \frac{\partial x_n}{\partial y_j}, \quad (8.10)$$

because $\partial x_k / \partial y_j = 0$ for $k \neq j, n$. In fact, it follows from the definition of \mathbf{y} , that $\partial x_j / \partial y_j = 1$ and $\partial x_n / \partial y_j = -1$. Thus,

$$\frac{\partial G_i}{\partial y_j} = \frac{\partial F_i}{\partial x_j} - \frac{\partial F_i}{\partial x_n}. \quad (8.11)$$

In matrix form, observe that

$$\begin{bmatrix} \frac{\partial G_1}{\partial y_1} & \dots & \frac{\partial G_1}{\partial y_{n-1}} \\ \vdots & \ddots & \vdots \\ \frac{\partial G_{n-1}}{\partial y_1} & \dots & \frac{\partial G_{n-1}}{\partial y_{n-1}} \end{bmatrix} = \begin{bmatrix} \frac{\partial F_1}{\partial x_1} & \dots & \frac{\partial F_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{n-1}}{\partial x_1} & \dots & \frac{\partial F_{n-1}}{\partial x_n} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{n-1} \\ -\mathbf{1}_{n-1}^\top \end{bmatrix}. \quad (8.12)$$

where \mathbf{I}_{n-1} and $\mathbf{1}_{n-1}$ are the $n-1$ dimensional identity matrix and column vector of all ones, respectively, as was defined in Section 2.1. The result showing the application of Theorem 8.3 to the DeGroot–Friedkin model is now ready to be presented.

Theorem 8.4. *Suppose that $\gamma_i < 1/2$ for all i . Then the map \mathbf{F} given in Eq. (8.7) has a unique fixed point in $\hat{\Delta}_n$, and this fixed point is locally exponentially stable.*

Proof. While one will need to use $d\mathbf{G}_y$, for convenience the proof begins by studying certain properties of $\frac{\partial F}{\partial x}$, which allow for easier delivery of specific conclusions in relation to $d\mathbf{G}_y$. In summary, it will be proved that at any fixed point $\bar{\mathbf{x}} \in \hat{\Delta}_n$, $\frac{\partial F}{\partial x}$ has a single eigenvalue at zero and all other eigenvalues are real, positive, and with magnitude less than one. It will then be shown that the eigenvalues of $d\mathbf{G}_y$ are the nonzero eigenvalues of $\frac{\partial F}{\partial x}$, which allows Theorem 8.3 to be used.

Denote an arbitrary fixed point of \mathbf{F} as $\bar{\mathbf{x}}$. Then, clearly $F_i(\bar{\mathbf{x}}) = \bar{x}_i$ for any i , and

⁶The Jacobian calculated in Eq. (8.8) and Eq. (8.9) can, however, be used for nonlinear contraction analysis, as in Chapters 5 and 6, since that method studies the differential dynamics.

it follows that

$$\left. \frac{\partial F_i}{\partial x_i} \right|_{\bar{x}} = \bar{x}_i \quad (8.13)$$

$$\left. \frac{\partial F_i}{\partial x_j} \right|_{\bar{x}} = -\frac{\bar{x}_i \bar{x}_j}{1 - \bar{x}_j} \quad (8.14)$$

As discussed below Eq. (8.7), $\bar{x}_i > 0$ for all i . Since $\bar{x} \in \hat{\Delta}_n$, one immediately concludes that the diagonal entries of $\left. \frac{\partial F}{\partial x} \right|_{\bar{x}}$ are strictly positive and the off-diagonal entries strictly negative. Moreover, one can readily verify using Eq. (8.8) and Eq. (8.9) that the column sum of $\left. \frac{\partial F}{\partial x} \right|_{\bar{x}}$ is equal to zero for every column⁷. In other words, $\left[\left. \frac{\partial F}{\partial x} \right|_{\bar{x}} \right]^\top$ is the Laplacian matrix associated with a strongly connected directed graph, which implies that $\left. \frac{\partial F}{\partial x} \right|_{\bar{x}}$ has a single eigenvalue at zero and all other eigenvalues have positive real part (see Appendix A.2 and Lemma A.8 for a definition of the Laplacian of a graph and its spectral properties). Next, it will be shown that the other eigenvalues are strictly real and less than one in magnitude.

Define $A = \text{diag}(1 - \bar{x}_i)$ as a diagonal matrix with the i^{th} diagonal entry being $1 - \bar{x}_i$. Since $\bar{x}_i \in \hat{\Delta}_n$, A is positive definite. The matrix $B = \left. \frac{\partial F}{\partial x} \right|_{\bar{x}} A$ is symmetric, with diagonal entry $b_{ii} = \bar{x}_i(1 - \bar{x}_i) > 0$ and off-diagonal entries $b_{ij} = -\bar{x}_i \bar{x}_j < 0$. Verify that, for any i , there holds

$$\sum_{j=1}^n b_{ij} = \bar{x}_i(1 - \bar{x}_i) - \bar{x}_i \sum_{j=1, j \neq i}^n \bar{x}_j = \bar{x}_i(1 - \bar{x}_i - \sum_{j=1, j \neq i}^n \bar{x}_j) \quad (8.15)$$

$$= 0, \quad (8.16)$$

where the last equality was obtained by using the fact that $\bar{x} \in \hat{\Delta}_n \Leftrightarrow \sum_{j=1}^n \bar{x}_j = 1 \Leftrightarrow 1 - \bar{x}_i = \sum_{j=1, j \neq i}^n \bar{x}_j$. In other words, the row sum of B is equal to zero for every row. It follows that B is the Laplacian matrix of an undirected, complete graph; B has a single eigenvalue at zero and all other eigenvalues are positive real (see Lemma A.8 and the comment immediately below). Using Lemma A.4, one thus concludes that $\left. \frac{\partial F}{\partial x} \right|_{\bar{x}} = (A^{-1}B)^\top$ has all real eigenvalues (because A^{-1} is positive definite and B is positive semidefinite). Notice that $\text{trace}\left(\left. \frac{\partial F}{\partial x} \right|_{\bar{x}}\right) = \sum_{i=1}^n \bar{x}_i(1 - \bar{x}_i) = 1 - \sum_{i=1}^n \bar{x}_i^2 = \sum_{j=1}^n \lambda_j \left(\left. \frac{\partial F}{\partial x} \right|_{\bar{x}}\right)$, where λ_j is an eigenvalue of $\left. \frac{\partial F}{\partial x} \right|_{\bar{x}}$. Since $n \geq 3$ and $\left. \frac{\partial F}{\partial x} \right|_{\bar{x}}$ has only a single zero eigenvalue, it follows that all other eigenvalues of $\left. \frac{\partial F}{\partial x} \right|_{\bar{x}}$ are positive and strictly less than one (and real).

Define the matrix

$$T = \begin{bmatrix} I_{n-1} & \mathbf{0}_{n-1} \\ -\mathbf{1}_{n-1}^\top & 1 \end{bmatrix}, \quad T^{-1} = \begin{bmatrix} I_{n-1} & \mathbf{0}_{n-1} \\ \mathbf{1}_{n-1}^\top & 1 \end{bmatrix}, \quad (8.17)$$

where $\mathbf{0}_{n-1}$ is the $n - 1$ dimensional vector of all zeros. It was established earlier that $\left. \frac{\partial F}{\partial x} \right|_{\bar{x}}$ has column sum equal to zero, i.e. $\mathbf{1}^\top \left. \frac{\partial F}{\partial x} \right|_{\bar{x}} = \mathbf{0}^\top$. Combining this column sum fact

⁷The fact that the column sum of $\left. \frac{\partial F}{\partial x} \right|_{\bar{x}}$ is equal to zero for every column is a consequence of the property that $F : \hat{\Delta}_n \Rightarrow \hat{\Delta}_n$. This was also noted in Chapter 5, in Lemma 5.2.

with Eq. (8.12), observe then, that

$$\begin{bmatrix} d\mathbf{G}_y & \frac{\partial F}{\partial x_n} \\ \mathbf{0}_{n-1}^\top & 0 \end{bmatrix} = \mathbf{T}^{-1} \frac{\partial F}{\partial \mathbf{x}} \mathbf{T} \quad (8.18)$$

where $\frac{\partial F}{\partial x_n}$ is a column vector with i^{th} element $\frac{\partial F_i}{\partial x_n}$. The similarity transform in Eq. (8.18) indicates that the matrix on the left of Eq. (8.18) has the same eigenvalues as $\frac{\partial F}{\partial \mathbf{x}}$, and since the matrix is block triangular, it follows that $d\mathbf{G}_y$ has the same nonzero eigenvalues as $\frac{\partial F}{\partial \mathbf{x}}$.

Since the above assumed that $\bar{\mathbf{x}}$ was an arbitrary fixed point it follows that all eigenvalues of $\frac{\partial F}{\partial \mathbf{x}}$ at any fixed point in $\hat{\Delta}_n$ are real and strictly less than one, which in turn implies that the eigenvalues of $d\mathbf{G}_y$, at any fixed point $\bar{\mathbf{y}} = [\bar{x}_1, \dots, \bar{x}_{n-1}]^\top$, are inside the unit circle. By Theorem 8.3, \mathbf{G} has a unique fixed point $\bar{\mathbf{y}}$ in $\hat{\Delta}_n$, and thus \mathbf{F} in Eq. (8.7) has a unique fixed point $\bar{\mathbf{x}}$ in $\hat{\Delta}_n$. \square

Remark 8.2. The fact that the eigenvalues of $d\mathbf{G}_y$ at a point in $\hat{\Delta}_n$ are a subset of those of $\frac{\partial F}{\partial \mathbf{x}}$ is no surprise. Because $\hat{\Delta}_n$ is invariant under \mathbf{F} , the translation of the affine space enclosing the set to define a linear space (including the origin) must have the property that this linear space is an invariant subspace for $\frac{\partial F}{\partial \mathbf{x}}$. As such, linear algebra indicates that the eigenvalues of $\frac{\partial F}{\partial \mathbf{x}}$ restricted to the invariant subspace are a subset of the full set of eigenvalues of $\frac{\partial F}{\partial \mathbf{x}}$. The above proof chose to give a more “explicit” proof of the relation, in the process identifying the eigenvalue of $\frac{\partial F}{\partial \mathbf{x}}$ that drops out in restricting to the invariant subspace.

Remark 8.3. Note that it is possible to prove $\frac{\partial F}{\partial \mathbf{x}}$, for all $\mathbf{x} \in \hat{\Delta}_n$, has strictly real eigenvalues with a single zero eigenvalue and all others positive. This property holds not only at the fixed point of \mathbf{F} . However, via simulations, it was observed that the eigenvalues of $\frac{\partial F}{\partial \mathbf{x}}$ can be greater than one (other than at a fixed point), and this typically occurred near the boundary of $\hat{\Delta}_n$. This motivated the introduction of a differential coordinate transform in Chapters 5 and 6 to show the transformed Jacobian had 1-norm strictly less than one in $\hat{\Delta}_n$; nonlinear contraction analysis was used to conclude exponential convergence to a unique fixed point $\bar{\mathbf{x}}$. It is not always assured that such a transform exists; the one proposed in Chapter 5 and the subsequent proof of the norm upper bound in Theorem 5.2 were nontrivial and not intuitive. In this chapter, the analysis has been greatly simplified by looking at the Jacobian at only the fixed points of \mathbf{F} (which were initially assumed to not be unique). However, this method guarantees only **local** convergence, in the sense that although there can be only one fixed point, the existence of trajectories which are not convergent to the fixed point but rather for example converge to an orbit is not precluded. Moreover, the technique of Theorem 8.3 cannot be easily extended to treat non-autonomous versions of Eq. (8.1), which in the DeGroot–Friedkin model, occurs when the social network topology is dynamic. For the non-autonomous case, Chapter 6 uses the techniques of nonlinear contraction to conclude there is a **unique limiting trajectory** $\bar{\mathbf{x}}(k)$ for the DeGroot–Friedkin model.

Remark 8.4. It should be stressed that a key desirable aspect of Theorem 8.3 is that in applying it, one needs only to evaluate the Jacobian $d\mathbf{F}_x$ at the fixed points of \mathbf{F} . In contrast, recall that a standard method to prove that $\mathbf{F} : X \rightarrow X$ has a unique fixed point $\bar{\mathbf{x}}$ and that

Eq. (8.1) converges exponentially fast to \bar{x} is via Banach's Fixed Point Theorem (assuming X compact). Specifically, one sufficient condition for F to be a contractive map would be to prove that, for some matrix norm $\|\cdot\|$, there holds $\|dF_x\| < \alpha$, $\forall x \in X$, where $\alpha < 1$, with a further assumption that X be convex [Khamisi and Kirk, 2011]. Thus, global properties rather than local (at fixed point) properties are required to generate the conclusion. The nonlinear contraction analysis in Chapter 5 is an extension of this via a differential transform.

The difficulty in proving a contractive property is acute for general nonlinear F . A nonlinear F results in dF_x being state-dependent. Consider two consecutive points of the trajectory of Eq. (8.1) with neither a fixed point, which are denoted $x_1 = x(k)$ and $x_2 = x(k+1)$, and suppose that $dF_x|_{x_1}$ and $dF_x|_{x_2}$ both have eigenvalues with magnitudes less than 1, i.e. assume that the eigenvalue restriction applies other than just at the fixed points. Then according to [Horn and Johnson, 2012, Lemma 5.6.10], there exist norms $\|\cdot\|'$ and $\|\cdot\|''$ such that $\|dF_x|_{x_1}\|' < 1$ and $\|dF_x|_{x_2}\|'' < 1$. However, it cannot be guaranteed that there exists a single norm $\|\cdot\|'''$ such that $\|dF_x|_{x_1}\|''' < 1$ and $\|dF_x|_{x_2}\|''' < 1$. In this chapter, there was no need to consider norms, nor was it necessary to study eigenvalue properties at all points; one only needed to consider the eigenvalues of dF_x at fixed points $\bar{x} = F(\bar{x})$ to simultaneously obtain a unique fixed point conclusion and local exponential convergence.

While Theorem 8.3 does not deliver a global convergence result, it may be useful when paired with other analysis methods, e.g. analysis using Lyapunov functions, particularly when other analysis methods may be unable to establish whether the fixed point is unique.

8.4 Conclusions

This chapter has used a recent nonlinear mapping convergence result to analyse the DeGroot–Friedkin model. More specifically, the result in [Anderson and Ye, 2018] used Lefschetz fixed point theory to show that a map applied to a positive invariant set (which also satisfied some further properties on compactness, convexity, and orientability) could be shown to have a unique, locally exponentially stable fixed point if the Jacobian evaluated at every fixed point had certain eigenvalue properties. This result was applied to the DeGroot–Friedkin system, illustrating how the map could be considered acting on an n -dimensional Euclidean space (for a network with n individuals) or on an $n - 1$ dimensional embedded space. The analysis is greatly simplified when compared to Chapter 5 and the original work [Jia et al., 2015], though the conclusion is local convergence rather than global. Nonetheless, this chapter illustrates the usefulness of the result in [Anderson and Ye, 2018] and its ease of application; the result may be used to bolster existing techniques, such as analysis using Lyapunov functions.

Part III

Opinion Dynamics with Interdependent Topics

Continuous-Time Opinion Dynamics with Interdependent Topics

Chapter Summary

This chapter considers a model in which a network of individuals discuss opinions on a *set of logically interdependent topics simultaneously*. Thus, the work is different to the EPO model (Part I), which assumes a single topic, and DeGroot–Friedkin model (Part II), which assumes multiple topics are discussed sequentially. Also, the model is in continuous-time, which is a departure from the previous discrete-time models considered. Key to this model is a “logic matrix” which describes the logical interdependence relations between the topics. Beginning with a network where no stubborn individuals exist, a necessary and sufficient condition is obtained for guaranteeing that the network reaches a consensus of opinions for each separate topic. The discrete-time counterpart has been studied to some degree by others, but to the author’s knowledge, this chapter is the first to study a continuous-time model for logically interdependent topic discussion. Surprisingly, different types of conclusions come from continuous time as opposed to discrete time. In fact, a key finding is that in continuous time, the condition for consensus couples the logic matrix with the network topology, whereas consensus in the discrete-time model is ensured if separate conditions on the logic matrix and network topology are satisfied. Last, stubborn individuals are considered, with some sufficient conditions obtained for the opinion to converge to a configuration of disagreement.

9.1 Introduction

This chapter considers a continuous-time model for a network of individuals who each have a belief system on the same set of topics, i.e. the individuals discuss opinions on a set of interdependent topics. Continuous-time models have been considered from the very beginnings of research on agent-based opinion dynamics;

shortly after John French Jr. first proposed the French–Harary–DeGroot model in [French Jr, 1956], Robert Abelson proposed a continuous-time counterpart in [Abelson, 1964]. Over the years, a number of models have been treated in both continuous- and discrete-time, perhaps with some minor differences. A notable example is the Friedkin–Johnsen model (discrete-time) [Friedkin and Johnsen, 1990] and the Taylor model (continuous-time) [Taylor, 1968], which consider the presence of stubborn individuals in the network. Similarly, the Altafini model has been considered in both continuous-time [Altafini, 2013; Proskurnikov et al., 2016; Hendrickx, 2014], and in discrete-time. [Xia et al., 2016a; Liu et al., 2017a; Hendrickx, 2014]. The distributed DeGroot–Friedkin model has been studied in continuous-time in [Chen et al., 2017] and in discrete-time in [Xu et al., 2015; Xia et al., 2016b]. Several reasons motivate study of continuous-time models. First, it is of interest to observe whether, *mutatis mutandis*, the conditions for ensuring certain behaviour of the opinion dynamical system are the same for continuous- and discrete-time models. It turns out that this is not always the case. For example, for the Abelson model, a strongly connected network will ensure that opinions reach a consensus, whereas in the DeGroot model, the network is required to be strongly connected *and aperiodic* in order for opinions to reach a consensus. If the network is periodic, the opinions will oscillate indefinitely. Second, given the same network structure, differences in the limiting behaviour may arise because in continuous-time, opinions typically evolve in a smooth manner, while opinion values can “jump” from one time instant to another in discrete-time; an example of differences can be seen in [Amelkin et al., 2017] vs [Liu et al., 2017b]. Third, the tools available to study continuous-time systems are generally more mature (at least in the systems and control community) when compared to discrete-time systems, especially for nonlinear systems.

In [Parsegov et al., 2017], an extension to the Friedkin–Johnsen model was proposed to study a network of individuals discussing their opinions on a set of interdependent topics simultaneously. The set of interdependent topics form a “belief system”, as termed by Converse in his now classical paper [Converse, 1964]. That is, an individual’s position on Topic *A* may influence his/her position on Topic *B* due to the individual’s view of constraints or relations between the two topics. In particular, the individual’s belief system may be used to internally adjust the individual’s opinions on the set of topics to eliminate cognitive inconsistencies in the opinions on the different, but related topics. In [Friedkin et al., 2016b], the same authors as in [Parsegov et al., 2017] use this model to study the belief systems of the American population on three topics regarding the 2003 Invasion of Iraq. The topics were *A*: “Iraq possesses Weapons of Mass Destruction”, *B*: “These Weapons of Mass Destruction are a danger to the Middle East and the world in general”, and *C*: “A preemptive invasion of Iraq is justified.” These topics are clearly interdependent, and [Friedkin et al., 2016b] showed that a change in the opinions of the US public on Topic *A* resulted in a shift in the opinions of the US public on Topic *B* and *C* due to this interdependence; there was majority support for the invasion in 2003 (i.e. Topic *C*) since most people believed Topic *A* to be true, but after 2005, when it was declared that there were no weapons of mass destruction (Topic *A*), the majority shifted

to withdraw support for the invasion (Topic C).

The model in this chapter may be considered as a continuous-time counterpart to the work of [Parsegov et al., 2017; Friedkin et al., 2016b]. Also, the proposed model may also be considered as an extension to the Taylor model [Taylor, 1968] (see also Section 1.2.5). Assuming that the matrix describing the logical interdependences between the topics (called the logic matrix) is homogeneous for all individuals, networks where there are no stubborn individuals, and where there are stubborn individuals, are both studied. First, focus is placed on networks with no stubborn individuals, and a necessary and sufficient condition is obtained for guaranteeing a consensus is reached. It turns out that this condition depends on *both* the logic matrix and the network topology, which is a distinct, and significant departure from the discrete-time model. For a given logic matrix, a sufficient (though certainly not necessary) condition on the network topology is obtained for ensuring consensus; it turns out that consensus is guaranteed if the strength of interpersonal interactions is sufficiently weak. From these findings, a reason for the difference in conditions for consensus between the continuous- and discrete-time models is conjectured. The oscillations of an individual's opinions because of, separately, the cognitive process for securing a consistent belief system and the interpersonal interactions for securing a consensus, can be much larger in the continuous-time model. As a result, the combined oscillations due to the cognitive process and the interpersonal interactions may create instability. For a class of logic matrices with sufficiently weak couplings, a class of networks with sufficiently weak interactions is identified which guarantee consensus. Following this, networks with stubborn individuals are studied, and several sufficient conditions involving the stubbornness of individuals are derived for ensuring that the opinions converge to a steady state of persistent disagreement.

9.1.1 Chapter Organization

The organisation of the chapter is as follows. Section 9.2 develops, explains, and justifies the opinion dynamics model, including the important logic matrix describing the interdependence between the topics. Sections 9.3 and 9.4 present convergence results for networks with no stubborn individuals and with stubborn individuals, respectively. Illustrative simulations are provided in Section 9.5, with concluding remarks given in Section 9.6.

9.2 Development of the Continuous-Time Model

In this section, a general opinion dynamics model is first presented where stubborn individuals may exist in the network. The model with no stubborn individuals, which is a specialisation, is also detailed. In the second part of this section, the matrix describing the logical interdependences between the topics (i.e. the matrix which forms the belief system) is explored in detail using an illustrative example. As a result, several constraints are identified on the matrix, arising from the natural properties present due to the socio-psychological context of the model.

9.2.1 Opinion Dynamics Model and Problem Statement

The general opinion updating model for each individual is provided, and a formal problem statement is presented. Let $\mathbf{x}_i(t) = [x_i^1(t), \dots, x_i^d(t)]^\top \in \mathbb{R}^d$ be the vector of opinion values¹ held by individual i , at time t , on d different topics. Where there is no confusion, the time argument t will be dropped. The opinions of individual i evolve according to the following differential equation

$$\dot{\mathbf{x}}_i(t) = \sum_{j \in \mathcal{N}_i} a_{ij} \mathbf{C} (\mathbf{x}_j(t) - \mathbf{x}_i(t)) + (\mathbf{C} - \mathbf{I}_d) \mathbf{x}_i(t) + b_i (\mathbf{x}_i(0) - \mathbf{x}_i(t)), \quad (9.1)$$

where a_{ij} is the $(i, j)^{th}$ entry of the adjacency matrix \mathcal{A} of the directed graph \mathcal{G} . Continuous-time models require a slightly different approach to describing the graph topology, including use of the Laplacian matrix \mathcal{L} of \mathcal{G} . The additional connectivity term “directed spanning tree”² used to describe \mathcal{G} , and the spectral properties of \mathcal{L} are given in Appendix A.2, which also introduces the graph theory required in this chapter. The matrix $\mathbf{C} \in \mathbb{R}^{d \times d}$ is the same for each individual $i \in \mathcal{I}$ and represents the logical interdependence/coupling between different topics. The scalar $b_i \geq 0$ is a measure of individual i ’s stubbornness, or attachment to initial opinion values $\mathbf{x}_i(0)$.

The properties of the constant matrix \mathbf{C} , and its interpretation, will be explored in detail in the next subsection. The update rule Eq. (9.1) with the special case $\mathbf{C} = \mathbf{I}_d$ is in fact the Taylor model, and has been well studied in a wide range of literature e.g. [Proskurnikov and Tempo, 2017; Taylor, 1968; Cao et al., 2012] (in effect d independently evolving consensus problems). If $b_i = 0 \forall i$, i.e. there are no stubborn individuals, then Eq. (9.1) reduces to

$$\dot{\mathbf{x}}_i(t) = \sum_{j \in \mathcal{N}_i} a_{ij} \mathbf{C} (\mathbf{x}_j(t) - \mathbf{x}_i(t)) + (\mathbf{C} - \mathbf{I}_d) \mathbf{x}_i(t). \quad (9.2)$$

The compact form of the dynamical system describing a network of n individuals’ opinions evolving according to Eq. (9.1) can be expressed as

$$\dot{\mathbf{x}} = (\mathbf{I}_n \otimes (\mathbf{C} - \mathbf{I}_d)) \mathbf{x} - (\mathcal{L} \otimes \mathbf{C}) \mathbf{x} + (\mathbf{B} \otimes \mathbf{I}_d) (\mathbf{x}(0) - \mathbf{x}), \quad (9.3)$$

where $\mathbf{x} = [\mathbf{x}_1^\top, \dots, \mathbf{x}_n^\top]^\top \in \mathbb{R}^{nd}$ is the stacked vector of all individual opinion vectors \mathbf{x}_i and \mathcal{L} is the Laplacian matrix³ associated with the graph \mathcal{G} . The diagonal matrix $\mathbf{B} = \text{diag}(b_i)$ encodes individuals’ stubbornness. One can rearrange to obtain

$$\dot{\mathbf{x}} = - \left[\mathbf{I}_{nd} + (\mathcal{L} - \mathbf{I}_n) \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{I}_d \right] \mathbf{x} + (\mathbf{B} \otimes \mathbf{I}_d) \mathbf{x}(0). \quad (9.4)$$

¹See Section 2.3.3 for a discussion of how an opinion may be represented as a real number.

²Some literature use other terms, e.g. rooted out-branching [Mesbahi and Egerstedt, 2010b] or directed rooted tree.

³By way of a brief summary, the Laplacian \mathcal{L} has a single eigenvalue at 0, and all other eigenvalues have positive real part if and only if \mathcal{G} has a directed spanning tree. With an eigenvalue of \mathcal{L} denoted as $\lambda_i(\mathcal{L})$, $i \in \mathcal{I}$, the eigenvalues can be ordered as $\lambda_1(\mathcal{L}) = 0$, and $\Re(\lambda_j(\mathcal{L})) > 0, j = 2, \dots, n$. Associated with $\lambda_1(\mathcal{L}) = 0$ is the right eigenvector $\mathbf{1}_n$ and left eigenvector $\gamma^\top \geq 0$ satisfying $\gamma^\top \mathbf{1}_n = 1$.

When the network has no stubborn individuals, i.e. when all individuals' evolve according to Eq. (9.2), the network has the following dynamics:

$$\dot{\mathbf{x}} = -\left(\mathbf{I}_{nd} + (\mathcal{L} - \mathbf{I}_n) \otimes \mathbf{C}\right)\mathbf{x}. \quad (9.5)$$

The problem considered in this chapter can be captured in the following objective.

Objective 9.1. *Let a social network be represented by a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$. Supposing that all individuals' opinions evolve according to Eq. (9.1), one seeks to determine (i) the topological constraints on the graph \mathcal{G} , (ii) the conditions on the matrix \mathbf{C} , and (iii) the conditions on \mathbf{B} , which guarantee that as $t \rightarrow \infty$, opinions converge to a limit, i.e. $\lim_{t \rightarrow \infty} \mathbf{x}_i(t)$ exists $\forall i \in \mathcal{I}$.*

A special case of the objective is consensus of opinions. It is said that a consensus of opinions has been reached if

$$\lim_{t \rightarrow \infty} \|\mathbf{x}_i(t) - \mathbf{x}_j(t)\|_2 = 0, \quad \forall i, j \in \mathcal{I}, \quad (9.6)$$

and it will be shown that consensus can occur when there are no stubborn individuals in the network, or if there exist stubborn individuals and the special case of $\mathbf{x}_i(0) = \mathbf{x}_j(0), \forall i, j$ occurs. Note that in the case of multiple topics, a consensus is said to be reached if the opinions of all individuals within each topic reach a consensus; the final consensus value for different topics in general will be different.

It will become apparent in the sequel that the problem is addressed by studying the marriage and influence of three different aspects of the opinion dynamical system outlined in Objective 9.1, viz. the network topology, logical interdependence, and individual stubbornness, described in Eq. (9.4) by \mathcal{L} , \mathbf{C} , and \mathbf{B} respectively⁴.

Remark 9.1. *In [Taylor, 1968], motivation and justification is given for a generalisation of Eq. (9.1) where individual i remains attached to several other static opinions, in addition to i 's initial opinion $\mathbf{x}_i(0)$. Suppose that individual i considers m different constant inputs $\mathbf{u}_{i,1}, \dots, \mathbf{u}_{i,m}$, such as the initial opinions of his/her neighbours, or constant information sources, e.g. politicians or the media. Then, the last term in Eq. (9.1) becomes $\sum_{k=1}^m b_{i,k}(\mathbf{u}_{i,k} - \mathbf{x}_i(t)) = \bar{b}_i(\bar{\mathbf{u}}_i - \mathbf{x}_i(t))$ where $\bar{b}_i = \sum_{k=1}^m b_{i,k}$ and $\bar{\mathbf{u}}_i = \frac{1}{\sum_{k=1}^m b_{i,k}} \sum_{k=1}^m b_{i,k} \mathbf{u}_{i,k}$ is the aggregate influence of all external sources. This chapter focuses on convergence and stability analysis, results which hold for both Eq. (9.1) and the generalised model discussed in this remark. Future research will focus on how the inputs $\mathbf{u}_{i,1}, \dots, \mathbf{u}_{i,m}$ affect the final opinion distribution, $\lim_{t \rightarrow \infty} \mathbf{x}(t)$.*

⁴Some existing literature on consensus and containment control study agents with dynamics $\dot{\mathbf{z}}_i = \mathbf{F}\mathbf{z}_i + \mathbf{G}\mathbf{u}_i$, with control input $\mathbf{u}_i = \mathbf{K} \sum_{j \in \mathcal{N}_i} a_{ij}(\mathbf{x}_j - \mathbf{x}_i)$. A typical result requires the control gain $\mathbf{K} = \mathbf{G}^\top \mathbf{P}$ where $\mathbf{P} = \mathbf{P}^\top > 0$ is the solution to some algebraic Riccati equation, see e.g. [Qin et al., 2015]. Lemma A.4 indicates that $\mathbf{G}\mathbf{G}^\top \mathbf{P}$ has nonnegative, real eigenvalues. In this paper, $\mathbf{C} - \mathbf{I}_d$ and \mathbf{C} in Eq. (9.1) replace \mathbf{F} and $\mathbf{G}\mathbf{G}^\top \mathbf{P}$, respectively. The matrix \mathbf{C} is allowed to have complex eigenvalues, which greatly increases the number of different couplings/interdependencies between topics describable by \mathbf{C} . Moreover, \mathbf{F} is also a function of \mathbf{C} . All this means conditions for stability are different to existing results.

9.2.2 Interdependent Topics and the Logic Matrix

In [Parsegov et al., 2017; Friedkin et al., 2016b], the authors call C the matrix of the *multi-issues dependence structure* (MiDS). In this chapter, the matrix C will be called the *logic matrix* because it encodes the logical coupling between issues. An example is now provided to motivate C and demonstrate its purpose in a person's cognitive process for handling logically interdependent topics. In doing so, an assumption on C is proposed which arises from studying the effect of C on an individual's opinion formation. The motivation and justification for this assumption was not detailed in [Parsegov et al., 2017].

Consider two topics being simultaneously discussed; Topic 1: North Korea has nuclear weapons capable of reaching the USA. Topic 2: The USA will use its arsenal to preemptively eliminate North Korea's nuclear strike ability. Let individual i 's opinion vector be $x_i = [x_i^1, x_i^2]^\top$. For topic 1, if x_i^1 is positive (respectively negative) then individual i believes (respectively does not believe) that North Korea has nuclear weapons capable of reaching the USA. For topic 2, if x_i^2 is positive (respectively negative) then individual i believes (respectively does not believe) that the USA will preemptively strike North Korea. One possible logic matrix is given by

$$C = \begin{bmatrix} 1 & 0 \\ 0.7 & 0.3 \end{bmatrix}, \quad (9.7)$$

which indicates that the individual's opinion on whether the USA will strike North Korea depends heavily on the individual's opinion on whether North Korea has nuclear weapons capable of reaching the USA. Note that while the above C is row-stochastic, in general it is not required that C be row-stochastic (though other constraints will apply).

To gain further insight into constraints on C , and by way of example, suppose that an individual i has no neighbours and is maximally open to influence, i.e. $\mathcal{N}_i = \{\emptyset\}$ and $b_i = 0$. Then Eq. (9.1) becomes

$$\dot{x}_i = Cx_i - x_i. \quad (9.8)$$

Here, the matrix C is the logic matrix, and $(C - I_d)x_i$ is the difference between individual i 's current opinion x_i and his/her opinions after taking into consideration the logical interdependencies of the discussed topics, Cx_i . In other words, Eq. (9.8) represents the cognitive process used by individual i to simultaneously adjust his/her opinions to ensure he/she has a consistent set of attitudes and beliefs [Gawronski and Strack, 2012; Converse, 1964].

Now, suppose that individual i has initial opinions $x_i(0) = [1, -0.5]^\top$, i.e. i learns that North Korea does have nuclear weapons and but initially believes with some certainty that the USA will not strike North Korea. Then Eq. (9.8) with C given in Eq. (9.7) indicates that $\lim_{t \rightarrow \infty} x_i(t) = [1, 1]$. In other words, individual i has an initial opinion against USA preemptively striking North Korea, but i 's belief that North Korea has nuclear weapons threatening the US creates an *inconsistency*. Individual

i 's cognitive process Eq. (9.8) is used to adjust his/her opinions until a consistent belief system is held, i.e. i believes the USA will preemptively strike North Korea because North Korea has nuclear weapons.

The fact that Eq. (9.8) represents a cognitive process implies some constraints must be placed on the matrix C ; human cognitive processes are not arbitrary. It is clear that if Eq. (9.8) is asymptotically stable then $\lim_{t \rightarrow \infty} x_i(t) = \mathbf{0}_d$, i.e. all opinion values converge to zero. This is a non-generic cognitive process, which will be assumed to not occur in this thesis. It is also not expected that the belief system is oscillatory, i.e. $x_i(t)$ does not enter a limit cycle. Moreover, it is reasonably required that the belief system not be unstable, i.e. $\lim_{t \rightarrow \infty} \|x(t)\|_\infty = \infty$. In order for Eq. (9.8) to satisfy these properties, the following assumption may be placed on the matrix C .

Assumption 9.1 (The matrix C). *The matrix $C \in \mathbb{R}^{d \times d}$ is such that $\lambda_1(C) = 1$ is a simple eigenvalue with an associated right (respectively left) eigenvector ζ (respectively ζ^\top), such that the eigenvectors satisfy $\zeta^\top \zeta = 1$. Moreover, for the other $d - 1$ eigenvalues $\lambda_k(C)$, there holds $\Re(\lambda_k(C)) < 1, \forall k \neq 1$.*

It is noted here that ζ is a nullvector of $C - I_d$, and if individual i has no neighbours and is not stubborn, i.e. has dynamics Eq. (9.8), then the final set of opinions $x_i(\infty)$ is in the span of ζ . It will become apparent that ζ also plays a role in determining the final set of opinions for the network of individuals. For convenience, the set of topics is indexed by $\mathcal{J} = \{1, \dots, d\}$.

Remark 9.2. *One may naturally ask whether, with $b_i > 0$, the cognitive process varies from Eq. (9.8) to become*

$$\dot{x}_i(t) = (C - I_d)x_i(t) + b_i C(x_i(0) - x_i(t)), \quad (9.9)$$

or the second summand has no logic matrix C , i.e.

$$\dot{x}_i(t) = (C - I_d)x_i(t) + b_i(x_i(0) - x_i(t)). \quad (9.10)$$

The latter is the proposed model, and is consistent with Eq. (9.1). Suppose that Assumption 9.1 holds. For large b_i , it is clear that Eq. (9.9) may become unstable (since under Assumption 9.1, C may have an eigenvalue with negative real part) but Eq. (9.10) remains stable. The author argues that in general, larger values of stubbornness in an individual should not create an unstable belief system, and therefore Eq. (9.10) represents a stubborn individual's cognitive process. Moreover, Appendix Section 9.7 shows that the Euler approximation of Eq. (9.1) (a specialisation of which yields Eq. (9.10) rather than Eq. (9.9)) give the model in [Parsegov et al., 2017].

9.3 Consensus for Individuals with No Stubbornness

This section will consider networks with no stubborn individuals. First, a necessary and sufficient condition is obtained for consensus of opinions. Following that, several

sufficient conditions (each different in some aspect) are obtained which consider either a class of logic matrices, or a class of network topologies.

9.3.1 Stability and Convergence to Consensus

The convergence result on the fundamental algorithm Eq. (9.2), where all individuals are maximally open to influence, i.e. not stubborn: $b_i = 0, \forall i \in \mathcal{I}$, is first presented. The key result is summarised in the following theorem.

Theorem 9.1. *Let \mathbf{C} , which satisfies Assumption 9.1, and $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ be given. Then for all initial conditions $\mathbf{x}(0)$, the social network, with each individuals' opinions evolving according to Eq. (9.2), reaches a consensus on all topics exponentially fast⁵ if and only if*

$$\Re((1 - \lambda_i(\mathcal{L}))\lambda_k(\mathbf{C})) < 1, \forall i = 2, \dots, n \text{ and } k \in \mathcal{J}, \quad (9.11)$$

where $\lambda_i(\mathcal{L})$ the eigenvalues of the Laplacian matrix \mathcal{L} , with $\lambda_1(\mathcal{L}) = 0$. Moreover, the solution satisfies

$$\lim_{t \rightarrow \infty} \mathbf{x}_i(t) = (\boldsymbol{\gamma} \otimes \boldsymbol{\xi})^\top \mathbf{x}(0) \boldsymbol{\zeta}, \forall i \in \mathcal{I}, \quad (9.12)$$

where $\boldsymbol{\gamma}^\top \in \mathbb{R}^{1 \times n}$ is the normalised left eigenvector of \mathcal{L} associated with the 0 eigenvalue satisfying $\boldsymbol{\gamma}^\top \mathbf{1}_n = 1$, and $\boldsymbol{\xi}^\top$ and $\boldsymbol{\zeta}$ are eigenvectors of \mathbf{C} defined in Assumption 9.1.

Proof. First, observe that Eq. (9.11) holds only if $\lambda_i(\mathcal{L}), i = 2, \dots, n$ are nonzero, which in turn holds if and only if the graph \mathcal{G} has a directed spanning tree (see Lemma A.8). That is, satisfying Eq. (9.11) implies that \mathcal{G} has a directed spanning tree⁶. We first establish the sufficiency of Eq. (9.11).

With all individuals maximally open to influence, i.e. $b_i = 0 \forall i \in \mathcal{I}$, the opinions $\mathbf{x}(t)$ of the social network evolve according to Eq. (9.5). Denote $\mathbf{M} = -\mathbf{I}_{nd} + (\mathbf{I}_n - \mathcal{L}) \otimes \mathbf{C}$. Clearly the i^{th} eigenvalue of \mathbf{M} is equal to $-1 + \lambda_i(\mathbf{A})$ where $\lambda_i(\mathbf{A})$ is the i^{th} eigenvalue of $\mathbf{A} = (\mathbf{I}_n - \mathcal{L}) \otimes \mathbf{C}$. The associated eigenvector is \mathbf{v}_i , where \mathbf{v}_i is the eigenvector of \mathbf{A} associated with $\lambda_i(\mathbf{A})$. One has that $\lambda_i(\mathbf{A}) = \mu_k \varphi_l$, where μ_k and φ_l are eigenvalues of $\mathbf{I}_n - \mathcal{L}$ and \mathbf{C} respectively, $k \in \mathcal{I}, l \in \mathcal{J}$ (see Lemma A.5). Then, one can verify that $\mathbf{v}_i = \mathbf{u}_k \otimes \mathbf{w}_l$ is an eigenvector of \mathbf{A} associated with $\lambda_i(\mathbf{A})$, where \mathbf{u}_k and \mathbf{w}_l are eigenvectors of $\mathbf{I}_n - \mathcal{L}$ and \mathbf{C} associated with μ_k and φ_l , respectively. According to Assumption 9.1, \mathbf{C} has a single eigenvalue at 1, which is denoted as φ_1 . If \mathcal{G} has a directed spanning tree, then \mathcal{L} has a single eigenvalue at zero, which implies $\mathbf{I}_n - \mathcal{L}$ has a single eigenvalue at 1. Denote this eigenvalue as μ_1 . Then clearly, $\lambda_1 = \mu_1 \varphi_1 = 1$ is an eigenvalue of \mathbf{A} with right eigenvector $\mathbf{v}_1 = \mathbf{1}_n \otimes \boldsymbol{\zeta}$.

⁵Convergence is of course exponential, since Eq. (9.5) is an LTI system.

⁶The requirement of a directed spanning tree is not surprising, and appears frequently in consensus problems for multi-agent systems [Ren and Cao, 2011]. If \mathcal{G} does not have a directed spanning tree then there is at least one individual i , or a group of individuals forming a closed and strongly connected subgraph (see Appendix A.2), who does not consider any other opinions. It follows that for generic initial conditions, individual i 's opinions (or the closed group of individuals' opinions) evolve separately, and will not reach a consensus with the opinions of any other individual k in the social network.

For $\lambda_i = \mu_1 \varphi_l, l = 2, \dots, d$, clearly $\lambda_i = \varphi_l$ has real part strictly less than 1, because Assumption 9.1 has that $\Re(\varphi_l) < 1$. For $\lambda_i = \mu_k \varphi_l$ where $k \in \{2, \dots, n\}, l \in \mathcal{J}$, if Eq. (9.11) is satisfied then λ_i has real part strictly less than 1. This in turn implies that $\Re(\lambda_i) < 1$ for all $i \neq 1$, which implies that all eigenvalues of M have strictly negative real part, except for a single eigenvalue at the origin, with associated right eigenvector $v_1 = \mathbf{1}_n \otimes \zeta$. From linear systems theory, one has that $x(t) = e^{Mt}x(0) = Pe^{Jt}P^{-1}x(0)$, where P is such that $J = P^{-1}MP$, and J is the Jordan canonical form of M , ordered such that

$$J = \begin{bmatrix} 0 & \mathbf{0}_{(nd-1)}^\top \\ \mathbf{0}_{(nd-1)} & \Delta \end{bmatrix}. \quad (9.13)$$

The $nd - 1$ nonzero diagonal entries of Δ are the stable eigenvalues of M . One then obtains $\lim_{t \rightarrow \infty} x(t) = p_1 q_1^\top x(0)$ where p_1 and q_1^\top are right and left eigenvectors of M associated with the single zero eigenvalue, satisfying $p_1^\top q_1 = 1$. The above analysis yielded $p_1 = \mathbf{1}_n \otimes \zeta$. One can easily verify that $q_1^\top = (\gamma \otimes \xi)^\top$ and thus $\lim_{t \rightarrow \infty} x = (\gamma \otimes \xi)^\top x(0)(\mathbf{1}_n \otimes \zeta)$. In other words, $\lim_{t \rightarrow \infty} x_i(t) = x_j(t) = \alpha \zeta$ as in Eq. (9.12), with $\alpha = (\gamma \otimes \xi)^\top x(0)$. Eq. (9.6) is satisfied. The sufficiency of Eq. (9.11) has thus been established.

It remains for the necessity of Condition 9.11 to be established. Suppose that Eq. (9.11) is not satisfied. Then there is some $\lambda_i = \mu_k \varphi_l, k \in \{2, \dots, n\}, l \in \mathcal{J}$ such that the eigenvalue of M , $-1 + \lambda_i$, is in the closed right half-plane. The system is either unstable, or $-1 + \lambda_i$ is on the imaginary axis (possibly at the origin). In the latter case either a) there are now at least 2 eigenvalues of M at the origin, or b) M has a pair of purely imaginary eigenvalues. Regarding a), the system is either unstable (there is a Jordan block in J associated with the eigenvalue 0, of size at least 2×2), or there are at least two 1×1 Jordan blocks associated with a zero eigenvalue. In the second possibility, consider for convenience the case where there are precisely 2 zero eigenvalues. Then, x converges exponentially fast to a subspace spanned by $\{v_1, v_i\}$ where v_i is an eigenvector of M associated with eigenvalue $\lambda_i = \mu_k \varphi_l = 0$, $k \neq 1$. Because $k \neq 1$, $v_i = u_k \otimes w_l$ cannot take the form $\mathbf{1}_n \otimes w_l$, for some $w_l \in \mathbb{R}^d$, which implies that consensus is not reached for generic initial conditions. Regarding b), denote one of the imaginary eigenvalues as $\lambda_i = \mu_k \varphi_l$, and recall that $k \neq 1$. Then, the system oscillates but not in consensus because, similar to the above arguments, v_i associated with the imaginary λ_i cannot take the form $\mathbf{1}_n \otimes w_l$. The proof is complete. \square

Remark 9.3. Theorem 9.1 provides a necessary and sufficient condition for the stability of the opinion dynamics system. Clearly, whether consensus is achieved depends on the combination of the network topology, as encoded by \mathcal{L} , and the logical interdependence as encoded by C . Given a C satisfying Assumption 9.1, two different graphs \mathcal{G}_1 and \mathcal{G}_2 may have different stability properties. For networks with no stubborn individuals, this is in direct contrast to the discrete-time result in [Parsegov et al., 2017], which establishes that consensus is reached if and only if C is regular, and either $\lim_{k \rightarrow \infty} C^k = \mathbf{0}_{n \times n}$ or W is fully regular⁷.

⁷A matrix A is regular if $\lim_{k \rightarrow \infty} A^k$ exists, and a row-stochastic matrix $B \in \mathbb{R}^{n \times n}$ is fully regular if

Here, \mathbf{W} is the influence matrix, i.e. the discrete-time counterpart to \mathcal{L} . Thus, in discrete-time, consensus is guaranteed if both \mathbf{C} and \mathbf{W} separately satisfy the required conditions, whereas in the continuous-time model Eq. (9.4), the conditions intertwine \mathbf{C} and \mathcal{L} , making prediction of conditions for consensus significantly more difficult. Naïve analysis of \mathcal{L} and $\mathbf{I} - \mathbf{C}$ separately could lead to a conclusion that a consensus is reached, when in fact the system is unstable; this is illustrated in the simulations in Section 9.5.

The author conjectures that the difference between the continuous- and discrete-time models is due to the different assumptions on the logic matrix and network topology. In discrete-time, \mathbf{W} is row-stochastic (i.e. $\|\mathbf{W}\|_\infty = 1$) and \mathbf{C} is regular (which implies $\|\mathbf{C}\|' \leq 1$ for some matrix norm $\|\cdot\|'$). In continuous-time, there are no restrictions on the size of the entries in \mathcal{A} (and consequently no restrictions on the size of the entries in \mathcal{L}), and \mathbf{C} only needs to satisfy Assumption 9.1. This means separately, the cognitive process Eq. (9.8) and the consensus process (which involves \mathcal{L}) can have much larger oscillations (i.e. the complex eigenvalues have large moduli) in continuous-time. When the two processes occur separately, there is no risk of instability, but when individuals discuss interdependent topics simultaneously, the two processes may combine to cause instability. Remark 9.5 below explores this in more detail.

It may be difficult to verify the conditions in Theorem 9.1 for complex networks because the precise values of eigenvalues of both \mathcal{L}, \mathbf{C} are needed. Two results are now presented on sufficient conditions which guarantee consensus *using limited information about the network and the logic structure*. The detailed motivation for this is discussed in Remark 9.5 below.

Corollary 9.1. *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ be given and suppose that \mathcal{G} has a directed spanning tree. Then for any given \mathbf{C} satisfying Assumption 9.1, there exists a graph with the same node and edge set as \mathcal{G} but with different edge weights, $\bar{\mathcal{G}} = \{\mathcal{V}, \mathcal{E}, \bar{\mathcal{A}}\}$, such that consensus of opinions is achieved using Eq. (9.2).*

Proof. Let \mathcal{L} be the Laplacian associated with the graph \mathcal{G} . One can easily verify that

$$\Re((1 - \lambda_i(\mathcal{L}))\lambda_k(\mathbf{C})) = d_k - y_i d_k \pm z_i e_k, \quad (9.14)$$

where, without loss of generality, $\lambda_i(\mathcal{L}) = y_i \pm z_i j$ and $\lambda_k(\mathbf{C}) = d_k \pm e_k j$ are complex conjugate eigenvalues of \mathcal{L} and \mathbf{C} respectively. Here, $z_i, e_k > 0$, and $y_i > 0$ for all $i \geq 2$ since \mathcal{G} has a directed spanning tree (see Lemma A.8). Because Assumption 9.1 holds, $d_k \leq 1$, for all k . For $i = 2, \dots, n$ and $k = 1, \dots, d$, it follows that $\Re((1 - \lambda_i(\mathcal{L}))\lambda_k(\mathbf{C})) < 1 \Leftrightarrow d_k - y_i d_k + z_i e_k < 1$. Define $\bar{\mathcal{A}} = \alpha \mathcal{A}$, where $\alpha > 0$ is a constant which adjusts every edge weight. Let $\bar{\mathcal{L}}$ be the Laplacian associated with $\bar{\mathcal{G}} = (\mathcal{V}, \mathcal{E}, \bar{\mathcal{A}})$. Observe that $\Re((1 - \lambda_i(\bar{\mathcal{L}}))\lambda_k(\mathbf{C})) = \Re((1 - \alpha \lambda_i(\mathcal{L}))\lambda_k(\mathbf{C}))$. It follows that consensus of opinions is achieved on the graph $\bar{\mathcal{G}}$ if and only if $d_k - \alpha(y_i d_k - z_i e_k) < 1$. Since $d_k \leq 1$ under Assumption 9.1, there always exists a sufficiently small α for which $d_k - \alpha(y_i d_k - z_i e_k) < 1$ holds, because $y_i > 0 \forall i \geq 2$. The proof is complete. \square

$\lim_{k \rightarrow \infty} \mathbf{B}^k = \mathbf{1}_n \mathbf{v}^\top$ for some $\mathbf{v} \in \mathbb{R}^n$ [Bullo et al., 2009].

The above corollary states that for any given C , there always exists a graph $\bar{\mathcal{G}}$ for which consensus of opinions can be achieved. The proof gives a simple way to scale the edge weights, as captured by \mathcal{A} , by the same constant α . It is however, not the only way to scale the edge weights. Below, a more complex scaling method is presented, which requires only limited knowledge of the network topology and the logic structure C . The reader is encouraged to revisit the Geršgorin disc theorem, viz. Theorem A.3, as it will be used here and later in several other instances in this chapter.

Corollary 9.2. *Let C , which satisfies Assumption 9.1, and $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ be given. Suppose that \mathcal{G} has a directed spanning tree. Then consensus of opinions is achieved if⁸, for all $k \in \mathcal{J}$*

$$\bar{l} < \min \left\{ \frac{|1 - |\lambda_k| \cos(\theta_k)| (1 + \cos(\theta_k))}{|\lambda_k| \sin^2(\theta_k)}, 0.5 \right\}, \quad (9.15)$$

where $|\lambda_k| = |\lambda_k(C)|$ and $\tan(\theta_k) = e_k/d_k$ with $\lambda_k(C) = d_k \pm e_k j$. Here, $\bar{l} = \max_{i \in \mathcal{I}} l_{ii}$, and $l_{ii} = \sum_{j \in \mathcal{N}_i} a_{ij}$ is the i^{th} diagonal entry of \mathcal{L} .

Proof. The system Eq. (9.5) reaches a consensus if and only if Statement 9.11 in Theorem 9.1 is satisfied. This is equivalent to ensuring that

$$d_k - y_i d_k + z_i e_k < 1 \quad (9.16)$$

where $\lambda_i(\mathcal{L}) = y_i \pm z_i j$ and $\lambda_k(C) = d_k \pm e_k j$ are any eigenvalue of \mathcal{L} and C , respectively, except for $\lambda_1(\mathcal{L}) = 0$ and $\lambda_1(C) = 1$. According to Assumption 9.1, $d_k < 1$.

The definition of \mathcal{L} implies that it has nonnegative diagonal entries and nonpositive off-diagonal entries, and moreover each row sums to 0 (see Appendix A.2). Moreover, \mathcal{L} has precisely one eigenvalue at 0 since \mathcal{G} contains a directed spanning tree (see Lemma A.8). Combining these observations with Theorem A.3, one concludes that every nonzero eigenvalue of \mathcal{L} is contained in the disc centred at \bar{l} , with radius \bar{l} . Denote this disc as $D_{\bar{l}}$. The fact that $\bar{l} < 0.5$ implies $y_i < 1$ (from Theorem A.3). Thus, $d_k - y_i d_k < 1$ because $d_k < 1$. If $\lambda_k(C)$ is real, i.e. $e_k = 0$, then Eq. (9.16) is satisfied. If all eigenvalues of C are real, then $\bar{l} < 0.5$ ensures stability.

Consider now $e_k > 0$ for some k . Observe that Eq. (9.16) is implied by $z_i^2 e_k^2 < (1 - d_k + y_i d_k)^2$, which is in turn implied by

$$\bar{z}_i^2 e_k^2 < (1 - d_k + y_i d_k)^2, \quad (9.17)$$

where $\bar{z}_i \geq z_i$ is such that $\beta_i = y_i + \bar{z}_i j$ is on the boundary of $D_{\bar{l}}$. Because β_i is on the boundary of $D_{\bar{l}}$, it satisfies $(y_i - \bar{l})^2 + \bar{z}_i^2 = \bar{l}^2$, which yields $\bar{z}_i^2 = -y_i^2 + 2y_i \bar{l}$. Substituting into Eq. (9.17) yields

$$(-y_i^2 + 2y_i \bar{l}) e_k^2 < (1 - d_k + y_i d_k)^2. \quad (9.18)$$

⁸As will be evident in the corollary proof, the right of Eq. (9.15) is well defined, since the left term approaches $(1 + |\lambda_k|)/2|\lambda_k| > 1/2$ as $\theta_k \rightarrow \pi$. This indicates that Eq. (9.15) holds for real eigenvalues of C , i.e. $\theta_k = \{0, \pi\}$.

Expanding and rearranging for \bar{l} yields $\bar{l} < \frac{1}{2y_i} \frac{(1-d_k)^2}{e_k^2} + \frac{d_k(1-d_k)}{e_k^2} + \frac{y_i}{2} \frac{(d_k^2 + e_k^2)}{e_k^2}$, or

$$\bar{l} < \frac{f_k(y_i)}{|\lambda_k|^2 \sin^2(\theta_k)}, \quad (9.19)$$

where

$$f_k(y_i) = \left[\frac{(1-ac)^2 + y_i^2 c^2 + 2y_i ac(1-ac)}{2y_i} \right], \quad (9.20)$$

with $a = \cos(\theta_k)$ and $c = |\lambda_k|$. Recall that $y_i > 0$. Calculations show that $\bar{y}_i = |1-ac|/c > 0$ is a unique minimum of $f_k(y_i)$ for $y_i \in (0, \infty)$. Since the right hand side of Eq. (9.20) is always strictly positive, it follows that Eq. (9.19) is implied by $\bar{l} < f_k(\bar{y}_i)/|\lambda_k|^2 \sin^2(\theta_k)$, which after some rearranging yields

$$\bar{l} < \frac{|1 - |\lambda_k| \cos(\theta_k)| (1 + \cos(\theta_k))}{|\lambda_k| \sin^2(\theta_k)}. \quad (9.21)$$

The proof is completed by noting that Eq. (9.21) must hold for all k to guarantee that Eq. (9.15) holds. Note that $|1 - |\lambda_k| \cos(\theta_k)| \neq 0$ because $|\lambda_k| \cos(\theta_k) = d_k < 1$.

Consider the scenario where $\mathbf{C}(\eta)$ varies smoothly as a function of some parameter $\eta \in [a, b]$, and for some $\kappa \in (a, b)$, $\lambda_p(\mathbf{C}(\kappa))$ has negative real part. Suppose further that $\lambda_p(\mathbf{C}(\eta))$ is real for $\eta \leq \kappa$, and is complex for $\eta > \kappa$. Then, $\lim_{\eta \rightarrow \kappa} \theta_p = \pi$. Notice that, separately, $\lim_{\theta_p \rightarrow \pi} 1 + \cos(\theta_k) = 0$ and $\lim_{\theta_p \rightarrow \pi} \sin^2(\theta_k) = 0$. As will now be shown, Eq. (9.15) continues to hold, i.e. is evaluable, as θ_p approaches π . Define $g(\theta_p) = |\lambda_p| \sin^2(\theta_p)$ and $h(\theta_p) = (1 - |\lambda_p| \cos(\theta_p))(1 + \cos(\theta_p))$. Denote $\lim_{\theta_p \rightarrow \pi^-}$ as the limit of θ_p approaching π from the left. Since $h(\theta_p), g(\theta_p)$ are continuous in θ_p , calculations using L'Hôpital's rule yield

$$\lim_{\theta_p \rightarrow \pi^-} \frac{h(\theta_p)}{g(\theta_p)} = \lim_{\theta_p \rightarrow \pi^-} \frac{h'(\theta_p)}{g'(\theta_p)} = \frac{1 + |\lambda_p|}{2|\lambda_p|}. \quad (9.22)$$

That is, the limit exists. This is consistent with Eq. (9.15) because $(1 + |\lambda_p|)/2|\lambda_p| > 1/2$ for $|\lambda_p| > 0$. \square

Remark 9.4. Corollary 9.2 is a stronger result than Corollary 9.1. Corollary 9.1 requires the scaling of every a_{ij} by the same constant $\alpha > 0$. Corollary 9.2 shows that a_{ij} only need to be adjusted for individual i if $l_{ii} = \sum_{j \in \mathcal{N}_i} a_{ij}$ exceeds the right hand side of Eq. (9.15). Moreover, for individual i , a_{ij} do not need to be scaled by the same constant for different j .

Remark 9.5. Checking the sufficiency condition in Corollary 9.2 requires only limited information concerning \mathcal{L} and \mathbf{C} . Furthermore, consider the first term on the right hand side of Eq. (9.15). One can rewrite this as $\frac{(1-d_k)(1+\cos(\theta_k))}{e_k \sin(\theta_k)}$. Sets of topics whose \mathbf{C} have large e_k and θ_k close to $\pi/2$ are associated with a cognitive process, described in Eq. (9.8), where the opinions for an isolated individual would oscillate heavily and rapidly before settling to a consistent belief system. Given such a \mathbf{C} , Corollaries 9.1 and 9.2 show it is always possible to reach a consensus if there is a sufficiently slow exchange of opinions (weights a_{ij} are small).

The cost of guaranteeing consensus is a slower speed of convergence. On the other hand, rapid discussions (arising from large a_{ij} values) can create large oscillations in the opinions. When topics are uncoupled, i.e. $\mathbf{C} = \mathbf{I}_d$, there is no risk of instability with large a_{ij} , provided the graph has a directed spanning tree. However, when the topics are coupled, large oscillations in both the cognitive process and in the consensus process may lead to a collapse in the discussions. This is the marriage of two aspects of the opinion dynamical system, viz. network topology and logical interdependence, as alluded to in Sections 9.1 and 9.2. This is of particular interest because \mathbf{C} relates to a cognitive process and therefore **cannot be easily changed for a given set of issues**. It also provides commentary on the relative time scale (or intensity) of the cognitive process, i.e. Eq. (9.8), and the interpersonal interactions arising from the network, when establishing conditions required for consensus. Corollary 9.2 provides a straightforward method to determine the class of \mathcal{L} which ensures consensus of opinions for a given \mathbf{C} .

9.3.2 Consensus for a Class of Logic Matrices

As discussed in Section 2.3.3, in many opinion dynamics problems, it is desirable to scale the opinions to be in some predefined interval, with $[-1, 1]$ being popular. With \mathbf{x} and x_i defined as in Section 9.2.1, define $\mathcal{R} = \{\mathbf{x} : x_i^k \in [-1, 1], \forall i \in \mathcal{I}, \forall k \in \mathcal{J}\}$. It will now be shown that, under the assumption given below, $\mathbf{x}(0) \in \mathcal{R} \Rightarrow \mathbf{x}(t) \in \mathcal{R}, \forall t \geq 0$ and Eq. (9.6) is satisfied.

Assumption 9.2. (a) The i^{th} diagonal entry of the Laplacian matrix \mathcal{L} , associated with $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$, satisfies $l_{ii} \leq 1, \forall i \in \mathcal{I}$. (b) The k^{th} diagonal of the logic matrix \mathbf{C} satisfies $c_{kk} > 0, \forall k \in \mathcal{J}$, and $\|\mathbf{C}\|_\infty = 1$.

It is clear that Assumption 9.2 places constraints, including, separately, the size of the allowable oscillations, on both the logic matrix \mathbf{C} and the network topology \mathcal{L} . This was alluded to in Remark 9.3, which showed that conditions for consensus were tied to both \mathbf{C} and \mathcal{L} , and not just one or the other. Note also that $c_{kk} > 0$ is a reasonable assumption as it simply implies that the k^{th} topic is positively dependent on itself. No restrictions are placed on the signs of the off-diagonal entries of \mathbf{C} , i.e. how two different topics are coupled. First, it will be shown that \mathcal{R} is an invariant set for the system Eq. (9.5), and then it will be shown that consensus is always reached under Assumptions 9.1 and 9.2.

Lemma 9.1. Suppose that Assumption 9.2 holds for \mathbf{C} and $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$. Suppose further that each individual's opinion changes according to Eq. (9.2). Then, if $\mathbf{x}(0) \in \mathcal{R}$, the opinion vector $\mathbf{x}(t) \in \mathcal{R}$ for all $t \geq 0$, where \mathcal{R} was defined above Assumption 9.2.

Proof. The proof is presented in Appendix Section 9.7. □

To conclude Section 9.3, it will be shown that all networks that contain a directed spanning tree and satisfy Assumptions 9.1 and 9.2 will reach a consensus of opinions.

Theorem 9.2. Suppose that Assumptions 9.1 and 9.2 hold for a given \mathbf{C} and $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$. Then, with each individual's opinions evolving according to Eq. (9.2), for all initial conditions,

the social network will reach a consensus value on all topics with an exponential convergence rate if and only if \mathcal{G} has a directed spanning tree. The consensus value is given in Eq. (9.12)

Proof. The necessity of the directed spanning tree was explained in the proof of Theorem 9.1. Before sufficiency is established, some properties will first be established regarding the eigenvalues of the matrix $\mathbf{M} = -\mathbf{I}_{nd} + (\mathbf{I}_n - \mathcal{L}) \otimes \mathbf{C}$. Consider a given $l \in \{1, \dots, nd\}$. The l^{th} diagonal entry of \mathbf{M} is $m_{ll} = -1 + (1 - \sum_{j \in \mathcal{N}_i} a_{ij})c_{kk}$ for some $i \in \mathcal{I}$ and $k \in \mathcal{J}$. The off-diagonal entries of the l^{th} row, m_{lj} , are given by $(1 - l_{ii})c_{kp}$ for all $q \in \mathcal{I}, p \in \mathcal{J}, p \neq k$, and $a_{iq}c_{kp}$ for all $q \in \mathcal{I}, p \in \mathcal{J}$. From Assumption 9.2, one has $0 < c_{kk} \leq 1$ and $\sum_{j \in \mathcal{N}_i} a_{ij} = l_{ii} \leq 1 \forall i \in \mathcal{I} \Rightarrow 0 \leq 1 - \sum_{j \in \mathcal{N}_i} a_{ij} \leq 1$. It follows that $m_{ll} \leq 0$ for all $l \in \{1, \dots, nd\}$. Turning to Geršgorin disc theory (see Theorem A.3), define $R_l(\mathbf{M}) = \sum_{j=1, j \neq l}^{nd} |m_{lj}|$, i.e. the sum of the absolute values of the off-diagonal entries of the l^{th} row of \mathbf{M} . One can then verify that

$$R_l(\mathbf{M}) = |1 - l_{ii}| \sum_{p=1, p \neq k}^d |c_{kp}| + \sum_{j \in \mathcal{N}_i} a_{ij} \hat{c}_k = (1 - l_{ii}) \sum_{p=1, p \neq k}^d |c_{kp}| + l_{ii} \hat{c}_k, \quad (9.23)$$

where $\hat{c}_k = \sum_{p=1}^d |c_{kp}|$ is the sum of the absolute values of the entries in the k^{th} row of \mathbf{C} . Note that $0 \leq 1 - l_{ii} \leq 1$. Thus,

$$\begin{aligned} m_{ll} + R_l(\mathbf{M}) &= -1 + (1 - l_{ii})c_{kk} + (1 - l_{ii}) \sum_{p=1, p \neq k}^d |c_{kp}| + l_{ii} \hat{c}_k \\ &= -1 + (1 - l_{ii}) \left(\sum_{p=1, p \neq k}^d |c_{kp}| + c_{kk} \right) + l_{ii} \hat{c}_k \end{aligned} \quad (9.24)$$

From Assumption 9.2, there holds $\|\mathbf{C}\|_\infty = 1$ and $c_{kk} > 0$, which implies that $\sum_{p=1, p \neq k}^d |c_{kp}| + c_{kk} = \hat{c}_k \leq 1$. It follows that

$$m_{ll} + R_l(\mathbf{M}) = -1 + \hat{c}_k \leq 0. \quad (9.25)$$

This implies that $m_{ll} \leq -R_l(\mathbf{M})$, and that this holds for all $l \in \{1, \dots, nd\}$. Thus, the Geršgorin discs of \mathbf{M} are all in the left half-plane. Specifically, the discs are either in the open left half-plane ($m_{ll} < -R_l(\mathbf{M})$) or touch the imaginary axis at the origin but do not enclose it ($m_{ll} = -R_l(\mathbf{M})$, including the possibility that $m_{ll} = 0$). This implies that the eigenvalues of \mathbf{M} either have strictly negative real part, or are equal to zero.

It will now be established that if \mathcal{G} has a directed spanning tree, and Assumption 9.2 holds, then consensus is achieved. As in the proof of Theorem 9.1, define $\mathbf{A} = (\mathbf{I}_n - \mathcal{L}) \otimes \mathbf{C}$, with eigenvalue and eigenvector pair $\lambda_i(\mathbf{A})$ and \mathbf{v}_i . Define μ_k , \mathbf{u}_k and φ_l , \mathbf{w}_l as an eigenvalue and eigenvector pair of $(\mathbf{I}_n - \mathcal{L})$ and \mathbf{C} , respectively, $k \in \mathcal{I}$ and $l \in \mathcal{J}$. As in the proof of Theorem 9.1, one concludes that if \mathcal{G} contains a directed spanning tree, then $\lambda_1(\mathbf{M}) = -1 + \lambda_1(\mathbf{A}) = 0$ is an eigenvalue of \mathbf{M} , with associated right eigenvector $\mathbf{v}_1 = \mathbf{1}_n \otimes \boldsymbol{\zeta}$. Above, it was established that the eigenvalues of \mathbf{M} either have strictly negative real part, or are equal to zero. Note also that $\mu_1 = 1 \Rightarrow \lambda_i(\mathbf{A}) = \mu_1 \varphi_l \neq 1, \forall l \in \{2, \dots, d\}$ because $\Re(\varphi_l) < 1$ according to

Assumption 9.1. This implies that proving $\lambda_i(A) = \mu_k \varphi_l \neq 1 \forall k \in \{2, \dots, n\}$ and $l \in \mathcal{J}$ is equivalent to satisfying Condition 9.11 in Theorem 9.1.

Because \mathcal{L} has a single zero eigenvalue and all other eigenvalues have positive real part, it follows from Assumption 9.2 that $\Re(\mu_k) < 1$ for $k \neq 1$. This implies that $\lambda_i(A) = \mu_k \varphi_1 \neq 1$, for all $k \neq 1$, since $\varphi_1 = 1$ according to Assumption 9.1. Consider now $k \in \{2, \dots, n\}$ and $l \in \{2, \dots, d\}$. Because $l_{ii} \leq 1$, it follows from Theorem A.3 that $|\mu_k| \leq 1$. Because $c_{ll} > 0, \forall l \in \mathcal{J}$, $\|C\|_\infty = 1$, and $\varphi_1 = 1$ is simple eigenvalue, then for $l \neq 1$, the l^{th} Geršgorin disc of C is situated at c_{ll} with radius $1 - c_{ll}$. It follows that $|\varphi_l| < 1$ for $l = \{2, \dots, d\}$. Thus $|\lambda_i| = |\mu_k \varphi_l| \leq |\mu_k| |\varphi_l| < 1$ for all $k \in \{2, \dots, n\}$ and $l \in \{2, \dots, d\}$. In other words, Condition 9.11 of Theorem 9.1 is satisfied. The final consensus value is computed as in the proof of Theorem 9.1. \square

9.4 Persistent Disagreement for Stubborn Individuals

The chapter now turns to the study of networks with stubborn individuals, i.e. $\exists i \in \mathcal{I} : b_i > 0$. The compact form of the network dynamics is given in Eq. (9.4), and the convergence of the opinions to a finite limit as $t \rightarrow \infty$, i.e. Objective 9.1, is ensured if

$$\Re [\lambda_i((I_n - \mathcal{L}) \otimes C - B \otimes I_d)] < 1, \quad (9.26)$$

where $\lambda_i(\cdot)$ is an eigenvalue of $(I_n - \mathcal{L}) \otimes C - B \otimes I_d$. In other words, this section will seek conditions which ensure the exponential stability of the unforced system

$$\dot{x} = -[I_{nd} + ((\mathcal{L} - I_n) \otimes C) + B \otimes I_d] x \quad (9.27)$$

to the origin $x = \mathbf{0}_{nd}$. If this exponential stability condition is satisfied, one then concludes that the final opinions will converge to

$$\lim_{t \rightarrow \infty} x(t) = [I_{nd} + (\mathcal{L} - I_n) \otimes C + B \otimes I_d]^{-1} (B \otimes I_d) x(0). \quad (9.28)$$

Note that if $x_i(0) = x_j(0), \forall i, j \in \mathcal{I}$, i.e. all individuals are initially at consensus, then clearly $\dot{x} = (I_n \otimes (C - I_d))x$ and $\lim_{t \rightarrow \infty} x_i(t) = \xi^\top x_i(0)\zeta$ for all $i \in \mathcal{I}$, where ξ^\top and ζ were given in Assumption 9.1. When the initial conditions are not equal, the opinions converge to a steady state of persistent disagreement if the system Eq. (9.4) is stable. In what follows, several different conditions are given on the stubborn individuals which ensures that the system Eq. (9.4) is stable, for social networks whose C matrix satisfies Assumption 9.1. Then, under Assumption 9.2, a general condition is obtained on the stubbornness of the individuals for ensuring stability.

For this section, the Laplacian matrix \mathcal{L} associated with \mathcal{G} is assumed to be expressed as

$$\mathcal{L} = \begin{bmatrix} \mathcal{L}_{11} & \mathbf{0}_{r \times (n-r)} \\ \mathcal{L}_{21} & \mathcal{L}_{22} \end{bmatrix} \quad (9.29)$$

where with $r \leq n$, $\mathcal{L}_{11} \in \mathbb{R}^{r \times r}$ is the Laplacian associated with a closed (no incoming edges) and strongly connected subgraph. Note that for any graph \mathcal{G} which has a

directed spanning tree there exists a reordering of the nodes so that its associated Laplacian is given in the form in Eq. (9.29) (see Appendix A.2).

9.4.1 Stability for Social Networks Under Assumption 9.1

In this subsection, three results are presented on the stability of the opinion dynamical system. As a summary, stability conditions cover three scenarios: (i) the individuals are all slightly stubborn, (ii) the individuals are all extremely stubborn, and (iii) the individuals are stubborn to approximately the same degree.

9.4.1.1 Individuals Are Slightly Stubborn

Theorem 9.3. *Suppose that the matrix C satisfies Assumption 9.1, and that the system Eq. (9.5) (i.e. without stubborn individuals) reaches a consensus. Suppose further that the graph \mathcal{G} has nodes ordered so that its Laplacian \mathcal{L} is expressed in the form given in Eq. (9.29). The opinion dynamics system Eq. (9.4) with stubborn individuals converges to Eq. (9.28) exponentially fast if $\exists j \in \{1, \dots, r\} : b_j > 0$, and $b_i \geq 0$ are sufficiently small, for all $i \in \mathcal{I}$.*

Proof. In the proof of Theorem 9.1, it was established that if consensus is reached for the system Eq. (9.5), then $M = -I_{nd} + ((I_n - \mathcal{L}) \otimes C)$ has a single eigenvalue at zero, which was denoted as $\lambda_1(M) = 0$. It was shown in the same proof that $\lambda_1(M) = 0$ has an associated left eigenvector $u_1 = \mathbf{1}_n \otimes \zeta$ and right eigenvector $v_1^\top = \gamma^\top \otimes \xi^\top$ where ξ^\top and ζ are given in Assumption 9.1, and γ^\top is the normalised left eigenvector of the 0 eigenvalue of \mathcal{L} (see Lemma A.8).

The exponential stability property of the system Eq. (9.27) will now be established. Define $\mathbf{b} = [b_1, \dots, b_n]^\top$ and $B(\mathbf{b}) = \text{diag}(b_i)$. Next, by defining

$$Z(\mathbf{b}) = -I_{nd} + ((I_n - \mathcal{L}) \otimes C) - B(\mathbf{b}) \otimes I_d, \quad (9.30)$$

observe that the system Eq. (9.27) is equivalent to $\dot{x} = Z(\mathbf{b})x$. Observe that, with $b_i \geq 0$ and \mathbf{e}_i being the i^{th} canonical unit vector (see Section 2.1), there holds

$$\frac{\partial}{\partial b_i} [Z(\mathbf{b})] = -\mathbf{e}_i \mathbf{e}_i^\top \otimes I_d \leq 0. \quad (9.31)$$

Recall that $\lambda_1(Z(\mathbf{0}_n)) = 0$ has associated left and right eigenvectors $u_1 = \gamma^\top \otimes \xi^\top$ and $v_1 = \mathbf{1}_n \otimes \zeta$. Note that $\gamma^\top \geq 0$, with $\gamma_i > 0, i \in \{1, \dots, r\}$ (see Lemma A.8). It then follows from Lemma A.7 that

$$\frac{\partial}{\partial b_i} [\lambda_1(Z(\mathbf{0}_n))] = -(\gamma^\top \otimes \xi^\top)(\mathbf{e}_i \mathbf{e}_i^\top \otimes I_d)(\mathbf{1}_n \otimes \zeta) = -\gamma_i \leq 0. \quad (9.32)$$

Thus, for sufficiently small $b_i \geq 0, i \in \mathcal{I}$, the gradient $\frac{\partial}{\partial b_i} [\lambda_1(Z(\mathbf{0}_n))]$ is nonpositive. Moreover, it was assumed that $\exists j \in \{1, \dots, r\} : b_j > 0$, and because $\gamma_j > 0$ for $j \in \{1, \dots, r\}$ (see above), it follows that $\frac{\partial}{\partial b_j} [\lambda_1(Z(\mathbf{0}_n))] = -\gamma_j < 0$. I.e., the eigenvalue $\lambda_1(Z(\mathbf{0})) = 0$ moves into the open left half-plane as b_j increases from 0. Thus, for

sufficiently small b_j and $b_i, i \in \{1, \dots, n\} \setminus j$, $\lambda_1(\mathbf{Z}(\mathbf{b}))$ is a stable eigenvalue. The other $nd - 1$ eigenvalues are continuous functions of \mathbf{b} and thus will remain in the open left-half plane for small b_i (the eigenvalues are already in the open left half-plane by virtue of the fact that the nonstubborn system $\dot{\mathbf{x}} = \mathbf{Z}(\mathbf{0}_n)\mathbf{x}$ is assumed to be convergent). Thus, $\mathbf{Z}(\mathbf{b})$ is nonsingular, and by recalling the definition of $\mathbf{Z}(\mathbf{b})$ given in Eq. (9.30), the proof is completed. \square

9.4.1.2 All Individuals Are Extremely Stubborn

Lemma 9.2. *Suppose that the matrix \mathbf{C} satisfies Assumption 9.1, and that \mathcal{G} contains a directed spanning tree. The opinion dynamics system Eq. (9.4) with stubborn individuals converges to Eq. (9.28) exponentially fast if for all $i \in \mathcal{I}$, $b_i > 0$ is sufficiently large.*

Proof. The proof is an application of Geršgorin's Disk Theorem. Examination of Eq. (9.27) shows that $b_i > 0$ does not change the radius of the associated $(i-1)d + 1, (i-1)d, \dots, (i-1)d + d$ Geršgorin disks, but does shift the centre of the disc further along the negative real axis towards $-\infty$. If every b_i is sufficiently large, then every Geršgorin disk will be strictly inside the open left half-plane, indicating that the system Eq. (9.27) is exponentially stable about the origin $\mathbf{x} = \mathbf{0}_{nd}$. This, along with the arguments preceding Eq. (9.28), completes the proof. \square

Remark 9.6. *Lemma 9.2 does not require the system with no stubborn individuals to reach a consensus. In other words, high stubbornness in all individuals ensures the opinion dynamics process is convergent even when the topic couplings are complex. However, this may be at the expense of reaching a consensus. Future work may study an adaptive $b_i(t)$, which might capture individuals who increase their stubbornness when they sense the opinion evolution process is becoming unstable (an undesirable social situation).*

9.4.1.3 Individuals Have Almost the Same Stubbornness

Lemma 9.3. *Suppose that the matrix \mathbf{C} satisfies Assumption 9.1. Suppose further that the system Eq. (9.5) (i.e. without stubborn individuals) is convergent. Then, for any $\alpha > 0$, the opinion dynamics system Eq. (9.4) with stubborn individuals converges to Eq. (9.28) exponentially fast if $b_i = \alpha + \epsilon_i, \forall i \in \mathcal{I}$ and $\epsilon_i \in \mathbb{R}$ is sufficiently small.*

Proof. Suppose first that $\epsilon_i = 0, \forall i$. If $b_i = \alpha, \forall i$, then Eq. (9.27) yields

$$\dot{\mathbf{x}} = [-(1 + \alpha)\mathbf{I}_{nd} - (\mathcal{L} - \mathbf{I}_n) \otimes \mathbf{C}]\mathbf{x} \quad (9.33)$$

which implies that the eigenvalues of $\mathbf{Z} = -(1 + \alpha)\mathbf{I}_{nd} - (\mathcal{L} - \mathbf{I}_n) \otimes \mathbf{C}$ are the eigenvalues of $-\mathbf{I}_{nd} - (\mathcal{L} - \mathbf{I}_n) \otimes \mathbf{C}$ (which are all in the open left half-plane except for one at the origin) shifted along the real axis by $-\alpha < 0$. In other words, the eigenvalues of \mathbf{Z} have real part strictly less than zero, which ensures the exponential stability of Eq. (9.33) and ensures that the final opinions converge to Eq. (9.28).

Next, consider $\epsilon_i \neq 0$ for some i . From the fact that the eigenvalues of \mathbf{Z} are continuous functions of ϵ_i , one concludes that for sufficiently small ϵ_i , all eigenvalues

of $\bar{\mathbf{Z}} = -(1 + \alpha)\mathbf{I}_{nd} - (\mathcal{L} - \mathbf{I}_n) \otimes \mathbf{C} - \sum_{i=1}^n \epsilon_i \mathbf{e}_i \mathbf{e}_i^\top \otimes \mathbf{I}_d$ will remain in the open left-half plane. In other words, for minor perturbations induced by ϵ_i , $\bar{\mathbf{Z}}$ remains stable. \square

9.4.2 Stability for Social Networks Under Assumption 9.2

A more general result is now established for networks with stubborn individuals whose logic matrix \mathbf{C} and graph \mathcal{G} satisfy Assumptions 9.1 and 9.2.

Theorem 9.4. *Suppose that the logic coupling matrix \mathbf{C} and social network \mathcal{G} satisfy Assumptions 9.1 and 9.2. Suppose further that \mathcal{G} has a directed spanning tree with nodes ordered so that its Laplacian \mathcal{L} is expressed in the form given in Eq. (9.29). If each individual's opinions evolve according to Eq. (9.1), and $b_i \geq 0, \forall i \in \mathcal{I}$ and $\exists j \in \{1, \dots, r\} : b_j > 0$, then, the opinion dynamics system Eq. (9.4) converges to Eq. (9.28) exponentially fast.*

Proof. Observe that Eq. (9.27) can be written as $\dot{\mathbf{x}} = \bar{\mathbf{M}}\mathbf{x}$ where $\bar{\mathbf{M}} = \mathbf{M} - \mathbf{B} \otimes \mathbf{I}_d$, and $\mathbf{M} = -\mathbf{I}_{nd} + (\mathbf{I}_n - \mathcal{L}) \otimes \mathbf{C}$ was defined in the proof of Theorem 9.1. In that same proof, it was shown that the Geršgorin discs of \mathbf{M} were in the closed left half-plane, and the discs could at most touch the origin, but not enclose it. This implied that \mathbf{M} has eigenvalues that either have strictly negative real part, or are at the origin. If $b_i \geq 0, \forall i \in \mathcal{I}$ then $-\mathbf{B} \otimes \mathbf{I}_d$ is a diagonal matrix with nonpositive diagonal entries. It follows that, for $i \in \mathcal{I}$, the $(i-1)d+1, \dots, (i-1)d+d$ Geršgorin discs of $\bar{\mathbf{M}}$ are the $(i-1)d+1, \dots, (i-1)d+d$ Geršgorin discs of \mathbf{M} , with the same radius, but shifted along the real axis to the left by $b_i \geq 0$. In other words, the Geršgorin discs of $\bar{\mathbf{M}}$ are in the closed left half-plane, and can at most touch the origin, but not enclose it. By proving that $\bar{\mathbf{M}}$ is invertible (as shall be done now), one immediately proves that all eigenvalues of $\bar{\mathbf{M}}$ have negative real part, i.e. are all stable.

To establish a contradiction, suppose that $\bar{\mathbf{M}} = -\mathbf{I}_{nd} + (\mathbf{I}_n - \mathcal{L}) \otimes \mathbf{C} - \mathbf{B} \otimes \mathbf{I}_d$ is not invertible. Then there exists a nonzero vector $\mathbf{x} \in \mathbb{R}^{nd}$ such that $\bar{\mathbf{M}}\mathbf{x} = \mathbf{0}_{nd}$. This implies that $((\mathbf{B} + \mathbf{I}_n) \otimes \mathbf{I}_d) \mathbf{x} = ((\mathbf{I}_n - \mathcal{L}) \otimes \mathbf{C}) \mathbf{x}$, or

$$\mathbf{x} = \left(((\mathbf{B} + \mathbf{I}_n)^{-1}(\mathbf{I}_n - \mathcal{L})) \otimes \mathbf{C} \right) \mathbf{x}, \quad (9.34)$$

with $\mathbf{B} + \mathbf{I}_n$ invertible because $b_i \geq 0$ for all $i \in \mathcal{I}$. Obviously, Eq. (9.34) holds if and only if the matrix $\mathbf{N} = ((\mathbf{B} + \mathbf{I}_n)^{-1}(\mathbf{I}_n - \mathcal{L})) \otimes \mathbf{C}$ has an eigenvalue at 1. Notice that the diagonal entries of $(\mathbf{B} + \mathbf{I}_n)^{-1}$ are $(b_i + 1)^{-1} \leq 1$ for all $i \in \mathcal{I}$. It follows that, for $i \in \mathcal{I}$, the $(i-1)d+1, \dots, (i-1)d+d$ rows of $(\mathbf{B} + \mathbf{I}_n)^{-1}(\mathbf{I}_n - \mathcal{L})$ are equal to the same rows of $\mathbf{I}_n - \mathcal{L}$ scaled by $(b_i + 1)^{-1} \leq 1$. In the proof of Theorem 9.2, it was established that the eigenvalues of $\mathbf{I}_n - \mathcal{L}$ are in the closed unit circle. It follows Geršgorin disk theory that the eigenvalues of $(\mathbf{B} + \mathbf{I}_n)^{-1}(\mathbf{I}_n - \mathcal{L})$ are in the closed unit circle (see Theorem A.3).

The fact that \mathbf{N} does not have an eigenvalue at 1 is now proved by contradiction. Denote the eigenvalues of \mathbf{N} , $(\mathbf{B} + \mathbf{I}_n)^{-1}(\mathbf{I}_n - \mathcal{L})$, and \mathbf{C} as ψ_i , $\bar{\mu}_k$, and φ_l with associated eigenvectors \mathbf{v}_i , \mathbf{u}_k , and \mathbf{w}_l , respectively. One has that $\psi_i = \bar{\mu}_k \varphi_l$, for $k \in \mathcal{I}$ and $l \in \mathcal{J}$ (see Lemma A.5). Note that, from Assumption 9.1, $\varphi_1 = 1$ is a simple eigenvalue of \mathbf{C} . Using the same arguments as in the last paragraph of the proof of

Theorem 9.2, one can establish that under Assumption 9.2, $\psi_i = \bar{\mu}_k \varphi_l \neq 1$ for $k \in \mathcal{I}$ and $j \in \{2, \dots, d\}$. Thus, N has an eigenvalue at 1 if and only if $\psi_i = \bar{\mu}_k \varphi_1 = 1$ for some $k \in \mathcal{I}$. This implies that $\exists k : \bar{\mu}_k = 1$, i.e. for some nonzero $\mathbf{r} \in \mathbb{R}^n$, there holds $(\mathbf{B} + \mathbf{I}_n)^{-1}(\mathbf{I}_n - \mathcal{L})\mathbf{r} = \mathbf{r}$ or equivalently $(\mathbf{B} + \mathcal{L})\mathbf{r} = \mathbf{0}_n$. In other words, $\mathbf{B} + \mathcal{L}$ must be singular if N has an eigenvalue at 1. With \mathcal{L} expressed as in Eq. (9.29), one can use [Qu, 2009, Corollary 4.33] to verify that $\mathcal{L}_{11} + \text{diag}[b_1, \dots, b_r]$ is nonsingular because $\exists j \in \{1, \dots, r\} : b_j > 0$. It remains to verify that $\mathcal{L}_{22} + \text{diag}[b_{r+1}, \dots, b_n]$ is nonsingular. For convenience, denote $\text{diag}[b_{r+1}, \dots, b_n] = \mathbf{B}_{22}$. Consider

$$\bar{\mathcal{L}} = \begin{bmatrix} 0 & \mathbf{0}_{n-1}^\top \\ \mathcal{L}_{21}\mathbf{1}_{n-r} - \mathbf{B}_{22}\mathbf{1}_{n-r} & \mathcal{L}_{22} + \mathbf{B}_{22} \end{bmatrix}. \quad (9.35)$$

One can verify that $\bar{\mathcal{L}}$ is a Laplacian matrix associated with a graph $\bar{\mathcal{G}}$ that has a directed spanning tree⁹. According to Lemma A.8, $\bar{\mathcal{L}}$ has a single zero eigenvalue. The block triangular form of $\bar{\mathcal{L}}$ then implies that $\mathcal{L}_{22} + \mathbf{B}_{22}$ is nonsingular. Thus, $\mathbf{B} + \mathcal{L}$ is nonsingular. It follows that N does not have an eigenvalue at 1, and thus $\bar{\mathbf{M}}$ does not have an eigenvalue at 0; $\bar{\mathbf{M}}$ is Hurwitz. Eq. (9.28) gives the final opinions. \square

Remark 9.7. In [Parsegov et al., 2017], individual i is said to be **oblivious** if $b_i = 0$ and $\nexists j \in \mathcal{I}$ such that $b_j > 0$ and there is directed path from v_j to v_i . That is, an oblivious individual is not stubborn, and is not influenced by a stubborn individual's opinion via a directed path. The sufficient condition involving b_i in Theorem 9.4 is equivalent to ensuring that there are no oblivious individuals in the network, because there is at least one non-oblivious individual in the closed and strongly connected subgraph \mathcal{G}_L with associated \mathcal{L}_{11} .

9.5 Simulations

Simulations are now presented to illustrate some of the key results of this chapter. The network \mathcal{G} has the Laplacian

$$\mathcal{L} = \begin{bmatrix} 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -0.8 & 1 & -0.2 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -0.4 & 1 & 0 & -0.6 & 0 \\ 0 & 0 & -0.2 & 0 & -0.8 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ -0.3 & -0.7 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \quad (9.36)$$

⁹In fact, the graph $\bar{\mathcal{G}}$ is the graph \mathcal{G}_F detailed below Lemma A.8 with an additional node, denoted v_0 , and edges going from v_0 to $v_j, j \in \{r+1, \dots, n\}$ as captured by the column vector $\mathcal{L}_{21}\mathbf{1}_{n-r} - \mathbf{B}_{22}\mathbf{1}_{n-r}$.

The coupling matrix is given by

$$C = \begin{bmatrix} 1 & 0 & 0 \\ -0.1 & 0.2 & 0.7 \\ 0.1 & -0.8 & 0.1 \end{bmatrix}. \quad (9.37)$$

This may correspond to the following set of truth statements. Topic 1: North Korea has nuclear weapons capable of reaching the USA. Topic 2: As its ally, China will defend North Korea against an attack. Topic 3: The USA will use its nuclear arsenal to eliminate North Korea's nuclear strike ability. For each topic, a positive (respectively negative) opinion would represent an individual's certainty that the statement of the topic is true (respectively certainty that the statement of the topic is false). Verify that \mathcal{G} contains a directed spanning tree, and is of the form in Eq. (9.29), with $r = 4$. The initial conditions are generated with each $x_i^k(0)$ selected from a uniform distribution in the interval $[-1, 1]$. Initially, it is assumed that $b_i = 0, \forall i \in \mathcal{I}$ (no stubbornness).

One can verify that \mathcal{L} and C satisfy Assumptions 9.1 and 9.2. As predicted by Theorems 9.1 and 9.2, one observes in Fig. 9.1 that the opinions for all 3 topics reach a consensus. In contrast, for the same graph \mathcal{G} with the same associated \mathcal{L} and $x(0)$, Fig. 9.2 shows the case where the topics are uncoupled with $C = I_d$, i.e. regular consensus. With C as in Eq. (9.37), Topic 2 is coupled to Topics 1 and 3 by a negative and positive weight respectively. The coupling effect is clear: the consensus value of Topic 2 in Fig. 9.1 is further from the consensus value of Topic 1 and closer to the consensus value of Topic 3 when compared to Fig. 9.2. In the social context, this simulation reflects a group of individuals who believe China is less willing to help a North Korea who is actively developing nuclear warheads (Topic 2's negative coupling to Topic 1), but also believe China would come to the aid of North Korea if the USA were to attack (Topic 2's positive coupling to Topic 3).

In Fig. 9.3, the same C , as in Eq. (9.37), and $x(0)$ are used but each edge weight is multiplied by 3, i.e. $\tilde{\mathcal{L}} = 3\mathcal{L}$, where \mathcal{L} is given in Eq. (9.36). Condition 9.11 of Theorem 9.1 is not satisfied, and thus the opinions diverge at an exponential rate. This illustrates a different effect of having logical interdependences, which contrasts with the above paragraph: consensus is guaranteed for $\tilde{\mathcal{L}}$ with uncoupled $C = I_d$: since both \mathcal{L} and $I_d - C$ are Hurwitz, naïve analysis of the system would have lead one to believe that the combined dynamics, i.e. Eq. (9.5), are convergent. Last, stubbornness is introduced, with $\mathbf{b} = [0, 0.1, 0, 0.05, 0, 0.4, 0, 0.3]^\top$, and the same $x(0)$ as above. The \mathcal{L} and C are given in Eq. (9.36) and Eq. (9.37). As predicted by Theorem 9.4, opinions converge to a steady-state configuration of persistent disagreement. Notice that stubbornness greatly changes the final opinion distribution for each topic (compare with Fig. 9.1), even though only half the individuals are stubborn.

9.6 Conclusions

This chapter studied a continuous-time opinion dynamics model that considers an influence network where individuals discuss multiple logically interdependent to-

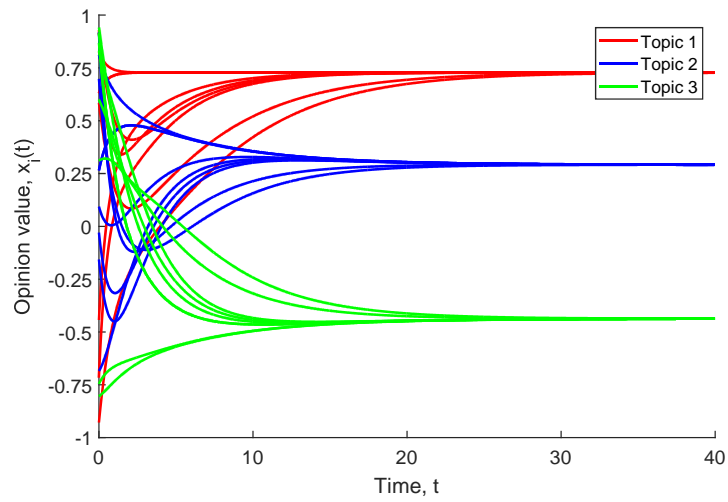


Figure 9.1: Consensus is reached on all 3 topics when the conditions of Theorem 9.1 are satisfied.

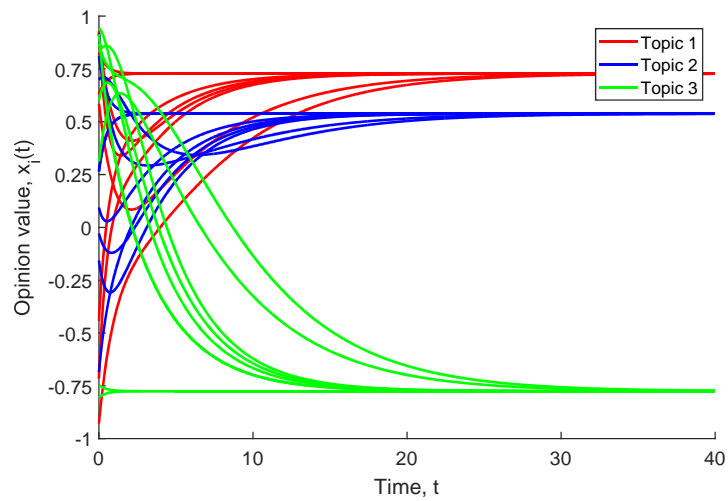


Figure 9.2: When the topics are uncoupled, $C = I_d$, consensus is reached but the final consensus values are different due to the lack of logic coupling.

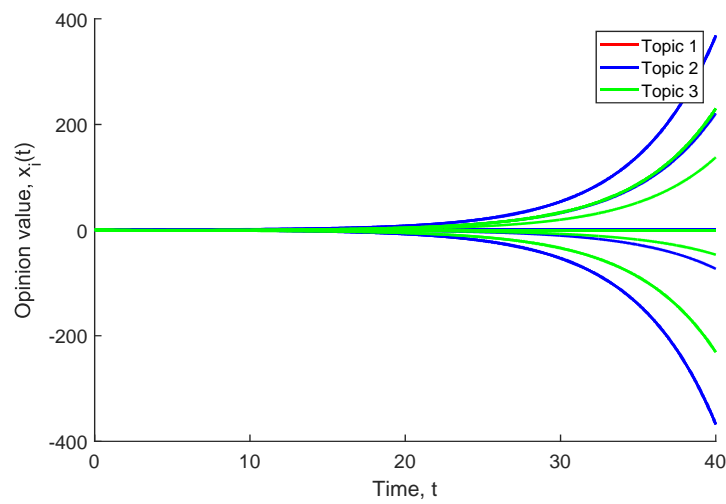


Figure 9.3: When Condition 9.11 of Theorem 9.1 is not satisfied, the opinion system is unstable.

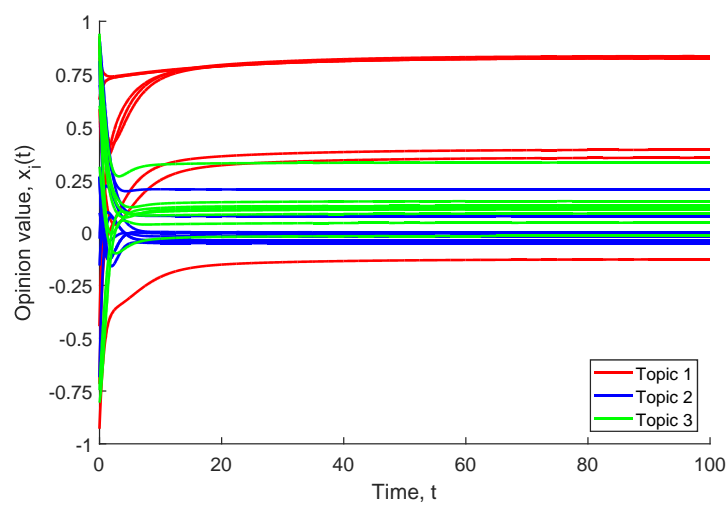


Figure 9.4: In the presence of four stubborn individuals, a state of persistent disagreement is achieved when the conditions in Theorem 9.4 are met.

pics simultaneously. After presentation of the model, the logic matrix describing the belief system for the interdependent topics was explored in detail. Beginning with networks that have no stubborn individuals, a necessary and sufficient condition was obtained for consensus to be reached. The condition depends on both the network topology and the logic matrix, which was a significant departure from the discrete-time model. It was conjectured that this difference is because in continuous-time, there are fewer restrictions on the logic matrix and network topology when each is viewed separately. Further exploration led to sufficient conditions for consensus, and the conclusion that one method of achieving a consensus was to ensure that the interpersonal interactions (encoded by the network) were weaker when compared to the coupling in the logical interdependences between topics. Then, a class of logic matrices and a class of topologies were identified where consensus is guaranteed as long as the network has a directed spanning tree. Turning then to individuals with stubbornness, several sufficient conditions for stability were identified, requiring the level of stubbornness to either be (i) small for all individuals, (ii) large for all individuals, or (iii) approximately equal for all individuals. Last, for a class of logic matrices and a class of network topologies, a general condition on the stubbornness of the individuals is obtained for ensuring stability of the opinion dynamical system.

From the investigations conducted in this chapter, it is evident that the introduction of interdependent topics result not only in significant differences in the limiting opinion values (assuming the opinions converge to a limit), but the coupling between topics has a major influence on the stability of the system itself.

9.7 Appendix: Some Proofs

9.7.1 Proof of Lemma 9.1

As above Eq. (9.1), define the k^{th} opinion of individual i as $x_i^k(t)$. To prove the lemma statement, one only needs to prove that, for all $k \in \mathcal{J}$ and $i \in \mathcal{I}$, there holds

$$\dot{x}_i^k(t) \leq 0, \quad \text{if } x_i^k(t) = 1 \quad (9.38)$$

$$\dot{x}_i^k(t) \geq 0, \quad \text{if } x_i^k(t) = -1 \quad (9.39)$$

for $\mathbf{x}(t) \in \mathcal{R}$. Denote the k^{th} row of \mathbf{C} as \mathbf{c}_k^\top . Dropping the t argument for clarity, one obtains from Eq. (9.2):

$$\dot{x}_i^k = \sum_{j \in \mathcal{N}_i} a_{ij} \mathbf{c}_k^\top (\mathbf{x}_j - \mathbf{x}_i) + \mathbf{c}_k^\top \mathbf{x}_i - x_i^k = \sum_{j \in \mathcal{N}_i} a_{ij} \sum_{l=1}^d c_{kl} (x_j^l - x_i^l) + \sum_{l=1}^d c_{kl} x_i^l - x_i^k \quad (9.40)$$

$$= \sum_{j \in \mathcal{N}_i} a_{ij} \sum_{l=1}^d c_{kl} x_j^l + (1 - l_{ii}) \sum_{l=1}^d c_{kl} x_i^l - x_i^k, \quad (9.41)$$

for any $k \in \mathcal{J}$ and $i \in \mathcal{I}$, with the last equality obtained by noting that $\sum_{j \in \mathcal{N}_i} a_{ij} = l_{ii}$. With $x_i^k = 1$ and $\mathbf{x}(t) \in \mathcal{R}$, it follows from Eq. (9.41) that Eq. (9.38) holds if

$$(1 - l_{ii})c_{kk} - 1 + (1 - l_{ii}) \sum_{l=1, l \neq k}^d c_{kl}x_i^l + \sum_{j \in \mathcal{N}_i} a_{ij} \sum_{l=1}^d c_{kl}x_j^l \leq 0. \quad (9.42)$$

Observe that Eq. (9.42) is implied by

$$(1 - l_{ii})c_{kk} - 1 + |1 - l_{ii}| \sum_{l=1, l \neq k}^d |c_{kl}| + l_{ii} \sum_{l=1}^d |c_{kl}| \leq 0 \quad (9.43)$$

because $|x_j^l| \leq 1, \forall j \in \mathcal{I}, l \in \mathcal{J}$ (including x_i^l) and $\sum_{j \in \mathcal{N}_i} a_{ij} = l_{ii}$. Since $l_{ii} \leq 1$ and $c_{kk} + \sum_{l=1, l \neq k}^d |c_{kl}| = \sum_{l=1}^d |c_{kl}| \leq 1$ under Assumption 9.2, Eq. (9.43) evaluates to be

$$(1 - l_{ii}) \left(c_{kk} + \sum_{l=1, l \neq k}^d |c_{kl}| \right) + l_{ii} \sum_{l=1}^d |c_{kl}| - 1 \leq 0. \quad (9.44)$$

It follows that for all $t \geq 0$, for any $k \in \mathcal{J}$ and $i \in \mathcal{I}$, the inequality in Eq. (9.44) holds. Because Eq. (9.44) holds, then Eq. (9.42) holds, and thus Eq. (9.38) holds. One can readily obtain a similar proof for Eq. (9.39). The proof is complete. \square

9.7.2 Relation to Discrete-Time Model

It will now be shown via an Euler approximation that the discrete-time counterpart to Eq. (9.1) is the multi-dimensional Friedkin-Johnsen algorithm studied in [Parsegov et al., 2017]. The discrete-time opinion model in [Parsegov et al., 2017] is given by

$$\mathbf{x}_i(k+1) = \lambda_i \sum_{j=1}^n w_{ij} \mathbf{C} \mathbf{x}_j(k) + (1 - \lambda_i) \mathbf{x}_i(0), \quad (9.45)$$

where $\lambda_i \in [0, 1]$ is an individual's susceptibility to influence, w_{ij} are the influence weights, and it is assumed that $\sum_{j=1}^n w_{ij} = 1$. Note that w_{ii} is not necessarily equal to zero¹⁰. One may compare the above with the Friedkin-Johnsen model in Section 2.3.2. Because $w_{ii} = 1 - \sum_{j \neq i}^n w_{ij}$, Eq. (9.45) is equal to

$$\mathbf{x}_i(k+1) - \mathbf{x}_i(k) = \lambda_i \left(1 - \sum_{j=1, j \neq i}^n w_{ij} \right) \mathbf{C} \mathbf{x}_i(k) + \lambda_i \sum_{j=1, j \neq i}^n w_{ij} \mathbf{C} \mathbf{x}_j(k) + (1 - \lambda_i) \mathbf{x}_i(0) - \mathbf{x}_i(k).$$

¹⁰In discrete-time, self-loops are often allowed ($w_{ii} > 0$) whereas most continuous-time algorithms assume no self-loops ($a_{ii} = 0, \forall i$).

Subtracting and adding $\lambda_i x_i(k)$ from the right hand side, and rearranging, yields

$$x_i(k+1) - x_i(k) = \lambda_i \sum_{j \neq i}^n w_{ij} C(x_j(k) - x_i(k)) + \lambda_i (C - I_d) x_i(k) + (1 - \lambda_i) [x_i(0) - x_i(k)].$$

Recall that the Euler approximation method states that for $\dot{y} = f(y)$ there holds, approximately, $y((k+1)\Delta) = y(k\Delta) + \Delta f(y(k\Delta)) \Rightarrow (y((k+1)\Delta) - y(k\Delta))/\Delta = f(y(k\Delta))$, where Δ is the sampling time. One observes that the above equation is simply the following difference equation

$$\begin{aligned} \frac{x_i((k+1)\Delta) - x_i(k\Delta)}{\Delta} &= \lambda_i \sum_{j \neq i}^n w_{ij} C(x_j(k\Delta) - x_i(k\Delta)) + \lambda_i (C - I_d) x_i(k\Delta) \\ &\quad + (1 - \lambda_i) [x_i(0) - x_i(k\Delta)] \end{aligned} \quad (9.46)$$

with $\Delta = 1$. As Δ becomes arbitrarily small in equation Eq. (9.46), Eq. (9.1) is recovered, where a_{ij} is the *instantaneous* influence weighting and b_i is the *instantaneous* stubbornness weighting. Note that Eq. (9.1) is not a precise equivalent of Eq. (9.45) because of the treatment of λ_i . Comparing the right hand sides of Eq. (9.46) and Eq. (9.1), it is clear that λ_i and w_{ij} have been combined and replaced with a_{ij} in the first summand. In the second summand, there is no λ_i . In Eq. (9.1), and in contrast to Eq. (9.46), $1 - \lambda_i \in [0, 1]$ in the third summand of Eq. (9.46) was replaced with $b_i \geq 0$ in Eq. (9.1). Note that no assumptions are placed on a_{ij} other than $a_{ij} \geq 0, \forall i, j$. This is different to the assumption that $0 \leq w_{ij} \leq 1$ and $\sum_{j=1}^n w_{ij} = 1$.

Conclusions and Future Work

By way of conclusion, the main contributions of this thesis are now summarised and discussed. Possible future research directions are also outlined in Section 10.2.

10.1 Thesis Summary and Contributions

This thesis has presented and studied three different models for social networks in three distinct parts; the EPO (Expressed–Private–Opinion) model, the DeGroot–Friedkin model, and a model for simultaneous discussion of interdependent topics. A number of different, complex social phenomena that arise through interpersonal interactions, and the conditions on network topology and parameters of each individual that give rise to said phenomena, have been identified.

A fundamental aspect of the thesis has been the careful consideration of the socio-psychological concepts that ground, and give meaning to, the models. Analysis has not just been presented on arbitrary dynamical systems. The problems investigated were chosen for their relevance when considering the context and motivations of the model, and interpretation of the results led to the drawing of insightful conclusions in the social context. To capture the numerous challenges of social network research, a variety of problems have been considered, ranging from model development, examination of established socio-psychological concepts using the model, to convergence on dynamic network topologies and network modification. A summary and discussion of the main results and contributions now follow.

10.1.1 How Differences in Private and Expressed Opinions Arise

Part I introduced the novel EPO (Expressed–Private–Opinion) model. The model was motivated by a range of social science literature studying the phenomena that (i) individuals can have different private and expressed opinions, and (ii) in many social situations, there is a pressure to conform to the group norm or standard. In particular, the model assumed that each individual had a private and an expressed opinion, and the expressed opinion was assumed to be the individual’s private opinion altered due to a pressure to conform to either the global or local public opinion.

After establishing a general stability result for a strongly connected social network, study shifted to interpreting and analysing the final opinion distribution, lea-

ding to the drawing of several interesting conclusions. It was shown that for general networks, the private and expressed opinions converge to a state of *strong diversity*, where persistent disagreement exists, and there are no clusters of individuals with the same opinion. Moreover, a pressure to conform led to preference falsification, and so in general, the *individuals were more willing to agree when expressing opinions*. Each individual in the network had *different private and expressed opinions* because of the combination of (i) a pressure to conform, (ii) stubborn attachment to his/her initial opinion, and (iii) the strong connectedness of the influence network. A method for *estimating the disagreement in the private opinions* was presented. Last, the *propagating effect* of resilience to conformity was established by proving that each individual's resilience affected every other individual's expressed opinion.

Following this, the EPO model was used to re-examine the seminal experimental works of Solomon Asch, and the phenomenon of pluralistic ignorance. Asch's experiments identified three broad types of individual reactions when an individual's belief in a fact was challenged by a unanimous majority. The EPO model was able to explain and capture all three types of reactions, which to the author's knowledge is a first for agent-based models (the Friedkin-Johnsen model was able to capture two, but could not distinguish between one of those two and the third). Interestingly, the EPO model was also able to explain why an increasing majority size did not necessarily translate to an increasing distortion of the individual's belief. Separately, stubborn extremists (zealots) were shown to be able to create pluralistic ignorance; the general population's private and expressed opinions converged to wildly different values. By studying both scale-free and small-world networks, and using both the local and global public opinion variant of the EPO model, it was identified that direct influence from zealots was the primary driver of pluralistic ignorance; indirect influence from the zealots was less impactful. Providing the global public opinion to, and increasing connectivity among, the general population are effective ways of drowning out (and thus minimising) the zealots' ability to create pluralistic ignorance.

10.1.2 Evolution of Individual Social Power

Part II studied the DeGroot-Friedkin model, which described the evolution of individuals' social powers (or self-confidences) in a network discussing a sequence of issues (also referred to as "topics"). Nonlinear contraction analysis was used to first study a network with constant relative interaction topology, and then a network with issue-varying relative interaction topology. It was established that the process of DeGroot opinion discussion combined with reflected self-appraisal of social power resulted in a "self-regulating" network. This is in the sense that no matter what value an individual's initial social power takes (which can be considered as the individual's estimate of his/her own power, or perceived power, before any discussion begins), that initial value is forgotten exponentially fast. For constant topologies, each individual's social power converges to a constant value, while for issue-varying/dynamic topologies, each individual's social power converges to a unique limiting trajectory. The constant value depends only on the constant topology structure, while for the

unique limiting trajectory, the individual's social power for any issue depends only on the sequence of dynamic topologies of the prior issues.

Further investigation of the model yielded additional conclusions: (i) the stability and instability of all equilibria where a single individual held all of the social power were characterised, (ii) an explicit upper bound on an individual's social power in the limit of the topic sequence, dependent only on his/her centrality in the network, and (iii) the convergence rate for a general class of network topologies was established.

In the model context, the star graph is a special class of topologies; star graphs are the only strongly connected graph in which general initial conditions on the individuals' social powers lead to a single individual (at the centre of the star) accumulating all of the social power in the limit of the topic sequence. This motivated Chapter 7, which studied several modifications of the star graph; for each modification, necessary and sufficient conditions on the strength of certain interpersonal interactions were established so that the centre individual no longer had the largest social power (one strategy considered two stars joined together, and the conditions for the two centre individuals to maintain social dominance). The unifying conclusion found through all the modifications was that for any individual wanting to change the social dominance of the network, particularly for autocratic networks, the individual may not be able to achieve the desired change by only modify relationships with his/her neighbours; often, modification of relationships in other parts of the network are necessary to create the desired change in social dominance.

To conclude Part II, a recent nonlinear mapping convergence result was applied to the DeGroot–Friedkin model with constant topology. In doing so, the advantages and disadvantages of the mapping result were identified; an advantage was that the analysis (and specifically the computation of the Jacobian) need only be conducted at all fixed points of the map, while the disadvantage was that this established only local exponential stability (though it does simultaneously prove that the map has a unique fixed point). Nonetheless, the nonlinear mapping result may prove useful as a complement to other analysis tools.

10.1.3 Opinion Dynamics with Interdependent Topics

Part III considered a continuous-time model where individuals discussed multiple logically interdependent topics. Introduction of the model included detailed explanation of how the logical interdependences between the topics could be captured by a "logic matrix". The properties of the logic matrix, arising from consideration of its function in an individual's cognitive process in establishing a consistent belief system on a set of topics, were explicitly established and justified.

Beginning with the assumption that the network contained no stubborn individuals, a necessary and sufficient condition was established for the opinions to reach a consensus. The condition involves both the logical interdependence and the network topology, clearly showing the added effects of the logical couplings. By further study and the establishing of a sufficient condition, a powerful conclusion can be drawn: for topic sets that have complex logical interdependences, consensus can be ensured

if the strength of the interpersonal influences is sufficiently weak. One interpretation is that interpersonal influences cannot “overpower” or have a greater impact than an individual’s own cognitive process in establishing a consistent belief system. From a time scale perspective, this indicates that the time scale associated with the consensus process (interpersonal interactions) must be *slower* than the time scale associated with an individual’s process of *securing logical consistency* in his/her belief system (the logic matrix).

The assumption that there are no stubborn individuals is relaxed, and several sufficient conditions for the stability of the opinion dynamical system is provided; the presence of stubborn individuals leads to a *strong diversity* of opinions, i.e. persistent disagreement. For networks where (i) the strength of the interpersonal influence is upper bounded, and (ii) the logic matrix has infinity norm equal to 1 (i.e. the logical coupling strength is upper bounded), consensus is shown to be guaranteed so long as the network contains a directed spanning tree. For the same class of networks satisfying the above two constraints, stability is shown to hold for a general condition on the stubbornness of the individuals in the network.

10.2 Future Work

In this section some future research problems are outlined. Throughout the chapters, a number of remarks have detailed potential research problems; no attempt is made to list them all, and only those of most interest (including some not detailed previously) are given below. The first three subsections, 10.2.1, 10.2.2, and 10.2.3 discuss near term future works on topics in Part I, II, and III, respectively. Subsection 10.2.4 discusses two long term problems of interest, both concerning the EPO model.

10.2.1 The EPO Model

A number of natural avenues for future research exist. The first is to provide a comprehensive theoretical treatment of the model by establishing stability properties of the model for (i) susceptibility and resilience values on the closed interval from 0 to 1 (whereas currently only values on the open interval are considered), and (ii) directed networks which are not strongly connected, e.g. networks which contain a directed spanning tree. This is of interest in order to verify the continuity of the model (which one expects to hold) as the parameters approach their upper and lower limits, with the primary difficulty arising because the result requires consideration of all three aspects (susceptibility, resilience, connectivity of the graph). Asynchronous, or gossip-based versions of this model are also of interest.

Second, the model may be made more complex, to reflect more complex human behaviour. A number of possibilities exist. One is to assume that the private and expressed opinion dynamics evolve at different time-scales. For example, an individual may only express an opinion, say at every 5 time steps, while his/her private opinion evolves at every time step. The resilience of an individual, captured by ϕ_i , was assumed to be constant in this thesis. There is evidence [Waters and Hans,

2009; Asch, 1951; Schachter, 1951] that ϕ_i may, among other possibilities, be state- and time-dependent, e.g. $\phi_i = \phi_i(y_i, \hat{y}_1, \dots, \hat{y}_n, t)$. The pressure may be proportional to the difference between an individual's private opinion and the public opinion, $|y_i - \hat{y}_{\text{avg}}|$ [Gorden, 1952; Schachter, 1951]. Pressure may also increase over time if an individual fails to conform to the public opinion [Waters and Hans, 2009]. Last, it is well known that large differences between an individual's private and expressed opinion generate discontent. If the difference is too large, an individual who initially conformed may become openly rebellious [Kuran, 1997]. This suggests ϕ_i might be a switching function activated by a state-dependent threshold.

The investigation on pluralistic ignorance can also be furthered in two aspects. The first is to study the potential behaviour when there are two sets of zealots at either end of the opinion spectrum. One might investigate the network or distributions of susceptibility and resilience parameters among the individuals that most commonly result in polarisation. This could indicate whether it is better to counteract false propaganda (the first set of zealots) with more false propaganda on the opposite of the opinion spectrum (the second set of zealots) or with the truth (e.g. more knowledge of the global public opinion). Moreover, theoretical research may establish the level of pluralistic ignorance given some limited knowledge of the network and the individuals' parameters. More generally, this research direction falls into the rapidly expanding area of network controllability, or *structural controllability* [Liu and Morse, 2017; Lin, 1974; Cowan et al., 2012]. Beyond the basic question of where to place zealots (controllable nodes) in order to drive the network to a given position, one may also investigate the types of strategies which would maximise the effect of zealots while minimising energy expenditure (measured by some metric such as number of zealots). For the all questions regarding controllability, the converse problems associated with increasing *robustness* of the network to malicious control may also be analysed; such questions have already arisen in the study of cyber-physical systems and are of equal interest in opinion dynamics.

It would be of great interest to conduct experimental research on the EPO model using human subjects, to validate the model, as done for e.g., the Friedkin–Johnsen model [Friedkin and Johnsen, 2011; Friedkin and Bullo, 2017]. This is a significant challenge that will require collaboration with social scientists: validation may serve as the ideal way to draw further interest from the sociology community to further develop the model in both the theoretical and experimental directions. The experimental methods detailed in [Friedkin and Johnsen, 2011] may provide a suitable starting point. One particular issue is the need to carefully design how an individual's private opinion is recorded: attempts to collect the private opinion may itself distort the private opinion value. A second issue is that two parameters, viz. resilience and stubbornness, must be determined for each individual.

10.2.2 The DeGroot–Friedkin Model

Future work on the DeGroot–Friedkin model includes a number of different paths. Remark 6.3 on page 108 proposed an interesting extension where individual behavi-

our, such as arrogance or humility, may be modelled into the self-appraisal process by way of a function applied to the true influence each individual has in the discussion of a given topic. A natural challenge is to identify the properties of the function which determine the limiting behaviour. In a second direction, the nonlinear contraction analysis technique might be applicable for study of the distributed DeGroot–Friedkin model, where opinion evolution and self-appraisal occur on the same time-scale; a recent result established exponential convergence for constant topology [Jia et al., 2017a], but it is unclear whether that work can be extended to dynamic topology.

Another problem of interest is to consider a state-dependent evolution of the relative interaction matrix, i.e. for topic s , one might have $C(x(0), \dots, x(s-1))$, to capture the fact that individuals may form/change interpersonal relationships based on social powers in the previous topic(s). One possibility is that for a given topic, individuals may seek to form a connection with the individual who had the greatest social power in the previous topic(s). In this case, initial (or perceived) social power may have a more profound effect on the limiting social power configuration. One key challenge will be to develop an update model for C that reflects how people develop or change interpersonal relationships, with support from social science literature.

For the work of Chapter 7, the next step is to generalise the strategies to consider an arbitrary number of attackers, dissenting subjects, etc. Beyond that, one may consider more arbitrary networks, or modification of topology to maximise (or minimise) the social power of a *group of individuals*. It is worth noting that the conditions obtained in Chapter 7 require only local information (strength of interpersonal relationships) about certain nodes. It would be of interest to see if, for arbitrary topologies, it is also possible to obtain conditions that require only local information, but result in global changes to the network properties (e.g. change of the socially dominant individual(s)).

10.2.3 Multiple Logically Interdependent Topics

In terms of the model introduced in Part III, several directions of research are available. First, for homogeneous logic matrices, a more general condition for stability should be established for networks which contain stubborn individuals; currently, a general condition is available only when restricted to a class of networks and a class of logic matrices.

Second, it is of great interest to consider heterogeneous logic matrices among the individuals. A recently accepted conference paper from the author [Ye et al., 2018], whose ideas were not incorporated in earlier parts of the thesis, studied networks where individuals had heterogeneous logic matrices, but in discrete-time. It was found that *competing logical interdependences*, where the logic matrix had the same pattern of zeros and nonzero entries but some off-diagonal entries had opposite (competing) signs for different individuals, resulted in persistent disagreement of opinions at equilibrium *even when there were no stubborn individuals*. Whether this phenomenon also exists in the continuous-time case is certainly of interest. As de-

tailed in Section 1.2.2, how strong diversity of opinions arises in strongly connected social networks is of great interest to social scientists, and while a number of different explanations have been provided (stubbornness, social distancing, or a desire to be unique), there have not been any works linking strong diversity to heterogeneous, and competing logic structures in belief system dynamics.

Third, it was mentioned in Remark 9.6 that there is interest in extending the model to consider an adaptive stubbornness $b_i(t)$, which might be state-dependent. This may reflect an individual who closes himself/herself off from influence when the combined interpersonal influence and cognitive processing of logical interdependencies threaten to create an unstable opinion system. However, one would need to take great care in proposing how $b_i(t)$ might evolve, since it must necessarily be supported by social science literature.

10.2.4 Longer-Term Research Problems: Investigation of New Social Phenomena

There are two related problems of particular interest, both of which arise from the EPO model, and there is certainly a possibility of having both in the same framework.

The first is to consider a second-order¹ model with private and expressed opinions. This arises from the concept of “the spiral of silence” [Noelle-Neumann, 1993; Taylor, 1982], which identified that an individual’s propensity to voice an opinion is linked not only to the current public opinion (i.e. the group average) but also the individual’s view of *how the public opinion is shifting*. In other words, people have an inherent inclination to *act based on predictions*, which leads to the natural extension of a second-order model. Study of such a model may lead to the identifying of whether the predictive nature of people can exacerbate certain social phenomena, such as pluralistic ignorance.

The second is to consider an event-based model for capturing when an individual decides to express an opinion (currently each individual expresses his/her opinion at every time step). It is not hard for the reader to imagine being in a situation where the reader only voices an opinion when (i) others’ opinions appear to be moving too far away from the reader’s current opinion, or (ii) others’ opinions are similar and so the reader is emboldened to voice support, and the reader stays silent in all other instances. It would certainly be of interest to see whether very different opinion trajectories arise from the two possibilities. Again, the author stresses a key difference with existing event-based multi-agent systems literature. In the multi-agent systems framework, the trigger function determining when an event occurs is typically *designed by the engineer* to drive the system to a global objective, e.g. consensus, with considerations for, e.g. energy saving by minimising the number of events. In the context of influence networks, the “trigger function” would be proposed to accura-

¹In multi-agent systems research, study of single-integrator agents (Abelson and Taylor models) naturally led to research on agents with double-integrator, linear, and then nonlinear dynamics, in an attempt to capture more realistic dynamics of autonomous agents. The motivation here is entirely different, and is driven by social concepts.

tely capture an individual's process in deciding whether or not to voice an opinion (perhaps using either of the above two examples). There is no "global objective", insofar as reaching a desired network state, though each individual may have his/her own objective. Furthermore, the individual may be interested in minimising the number events, because expressing opinions too frequently may annoy others in the group, or reduce the ability of the individual to influence others when voicing an opinion in the future.

Appendix

This appendix contains some standard results that are used in various parts of the thesis.

A.1 Linear Algebra

Theorem A.1 ([Berman and Plemmons, 1979]). *Consider a nonnegative $n \times n$ matrix A . The following facts hold*

1. *The matrix A has an eigenvalue $\lambda_1(A) = \rho(A)$.*
2. *Associated with $\lambda_1(A)$ are a left and right eigenvector $\mathbf{u}^\top \geq 0$ and $\mathbf{v} \geq 0$, respectively.*

Theorem A.2 (Perron–Frobenius Theorem,[Berman and Plemmons, 1979]). *Consider a nonnegative $n \times n$ matrix A , and suppose that A is irreducible.*

1. *If $A > 0$, then, $\rho(A)$ is a simple eigenvalue, greater than the magnitude of any other eigenvalue.*
2. *If $A \geq 0$, then, $\lambda_1(A) = \rho(A)$ is a simple eigenvalue with an associated with left and right eigenvector $\mathbf{u}^\top > 0$ and $\mathbf{v} > 0$. Any eigenvalue of modulus $\rho(A)$ is also simple. Moreover, the number h of eigenvalues of A of modulus $\rho(A)$ is equal to the periodicity of the graph $\mathcal{G}[A]$.*

Lemma A.1. *Let $A \in \mathbb{R}^{n \times n}$ be a given square row-substochastic matrix. Suppose A has at least one row with row sum strictly less than one, i.e. $\exists i : \sum_{j=1}^n a_{ij} < 1$. Suppose further that A is irreducible. Then, $\rho(A) < 1$.*

Proof. This lemma is an immediate consequence of [Varga, 2009, Lemma 2.8]. □

Lemma A.2 ([Bernstein, 2009, pg. 108–109]). *Suppose that an invertible matrix $Z \in \mathbb{R}^{n \times n}$ is partitioned as*

$$Z = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \tag{A.1}$$

Suppose further that A, D are invertible. Then $D - CA^{-1}B$ and $A - BD^{-1}C$ are invertible, and the inverse of Z is given as

$$Z^{-1} = \begin{bmatrix} (A - BD^{-1}C)^{-1} & -A^{-1}B(D - CA^{-1}B)^{-1} \\ -D^{-1}C(A - BD^{-1}C)^{-1} & (D - CA^{-1}B)^{-1} \end{bmatrix} \quad (\text{A.2})$$

Lemma A.3 ([Bernstein, 2009, Fact 10.11.20]). Consider a matrix function $A(s) \in \mathbb{R}^{n \times n}$, where $s \in \mathcal{D} \subseteq \mathbb{R}$. Suppose that $A(s)$ is continuously differentiable, and invertible, for all $s \in \mathcal{D}$. Then

$$\frac{dA^{-1}(s)}{ds} = -A^{-1}(s) \left(\frac{dA(s)}{ds} \right) A^{-1}(s). \quad (\text{A.3})$$

Lemma A.4 ([Horn and Johnson, 2012, Corollary 7.6.2]). Let $A, B \in \mathbb{R}^{n \times n}$ be symmetric. If A is positive definite, then AB is diagonalizable and has real eigenvalues. If, in addition, B is positive definite or positive semidefinite, then the eigenvalues of AB are all strictly positive or nonnegative, respectively.

Lemma A.5 ([Bernstein, 2009, Proposition 7.1.10]). Let $\lambda_i, i \in \{1, \dots, n\}$ and $\mu_j, j \in \{1, \dots, d\}$ be an eigenvalue of $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{d \times d}$ respectively. Then the nd eigenvalues of $A \otimes B$ are given by $\tau_k = \lambda_i \mu_j$ for $i \in \{1, \dots, n\}$ and $\mu_j, j \in \{1, \dots, d\}$. Suppose further that u_i and v_j are eigenvectors of A and B associated with λ_i and μ_j , respectively. Then, $w_k = u_i \otimes v_j$ is an eigenvector of $A \otimes B$.

Lemma A.6 ([Horn and Johnson, 2012]). Suppose that $A \in \mathbb{R}^{n \times n}$ has a simple, real eigenvalue λ . Let u and v be, respectively, a left and right eigenvector of A associated with the eigenvalue λ , that is, $Av = \lambda v$ and $\lambda u^\top = u^\top A$. Then, u, v are real, and can be normalised to satisfy $u^\top v = 1$.

Theorem A.3 ([Horn and Johnson, 2012, Theorem 6.1.1]). Let $A = \{a_{ij}\} \in \mathbb{R}^{n \times n}$, and let $R_i(A) = \sum_{j=1, j \neq i}^n |a_{ij}|$ for $i = 1, \dots, n$ denote the absolute row sums of the off-diagonal elements of A . Consider the n Geršgorin discs

$$\{z \in \mathbb{C} : |z - a_{ii}| \leq R_i(A)\}, \quad i = 1, \dots, n \quad (\text{A.4})$$

The eigenvalues of A are in the union of the Geršgorin discs

$$H(A) = \bigcup_{i=1}^n \{z \in \mathbb{C} : |z - a_{ii}| \leq R_i(A)\} \quad (\text{A.5})$$

Furthermore, if the union of k of the n discs that comprise $H(A)$ forms a set $H_k(A)$ that is disjoint from the remaining $n - k$ discs, then $H_k(A)$ contains exactly k eigenvalues of A , counted according to their algebraic multiplicities.

Lemma A.7. Suppose that a matrix $Z(b)$ is a function of $b \in \mathbb{R}$, and that for \bar{b} , $Z(\bar{b})$ has a simple eigenvalue $\lambda(\bar{b})$ with associated left and right eigenvectors u^\top and v , respectively, normalised to satisfy $u^\top v = 1$. Then,

$$\left. \frac{d\lambda(b)}{db} \right|_{\bar{b}} = u^\top \left. \frac{dZ(b)}{db} \right|_{\bar{b}} v. \quad (\text{A.6})$$

Proof. Observe that there holds $\frac{d\mathbf{Z}(b)\mathbf{v}}{db}\big|_{\bar{b}} = \frac{d\lambda(\mathbf{Z}(b))\mathbf{v}}{db}\big|_{\bar{b}}$ which is equivalent to

$$\frac{d\mathbf{Z}(b)}{db}\big|_{\bar{b}}\mathbf{v} + \mathbf{Z}(b)\frac{d\mathbf{v}}{db}\big|_{\bar{b}} = \frac{d\lambda(\mathbf{Z}(b))}{db}\big|_{\bar{b}}\mathbf{v} + \lambda(\mathbf{Z}(b))\frac{d\mathbf{v}}{db}\big|_{\bar{b}}.$$

Premultiplying both sides of the above equation by \mathbf{u}^\top yields

$$\mathbf{u}^\top \frac{d\mathbf{Z}(b)}{db}\big|_{\bar{b}}\mathbf{v} + \mathbf{u}^\top \mathbf{Z}(b)\frac{d\mathbf{v}}{db}\big|_{\bar{b}} = \mathbf{u}^\top \frac{d\lambda(\mathbf{Z}(b))}{db}\big|_{\bar{b}}\mathbf{v} + \lambda(\mathbf{Z}(b))\mathbf{u}^\top \frac{d\mathbf{v}}{db}\big|_{\bar{b}}. \quad (\text{A.7})$$

By recalling that $\mathbf{u}^\top \mathbf{Z}(b) = \lambda(\mathbf{Z}(b))\mathbf{u}^\top$, it follows that the second term on the left hand side cancels the second term on the right hand side of the above. Additionally, $\mathbf{u}^\top \frac{d\lambda(\mathbf{Z}(b))}{db}\big|_{\bar{b}}\mathbf{v} = \frac{d\lambda(\mathbf{Z}(b))}{db}\big|_{\bar{b}}$ because $\mathbf{u}^\top \mathbf{v} = 1$. This yields Eq. (A.6). \square

A.1.1 Consensus Level and Coefficient of Ergodicity

This section will recall a common performance function and a coefficient of ergodicity for a stochastic matrix, for details see [Seneta, 2006, pg. 82]. For a vector $\mathbf{x} \in \mathbb{R}^n$, define the performance function $V(\mathbf{x}) : \mathbb{R}^n \mapsto \mathbb{R}$ as

$$V(\mathbf{x}) = \max_{i \in \{1, \dots, n\}} x_i - \min_{j \in \{1, \dots, n\}} x_j, \quad (\text{A.8})$$

In the context of this thesis, $V(\mathbf{y})$ may be used to measure the “level of consensus” in the vector of opinions $\mathbf{y}(t)$ of a network with n individuals. Notice in particular that consensus, as defined in Definition 2.3, is reached if and only if $V(\mathbf{y}) = 0$. Next consider the following coefficient of ergodicity, $\tau(\mathbf{A})$ for a square stochastic matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, defined as

$$\tau(\mathbf{A}) = \frac{1}{2} \max_{i, j \in \{1, \dots, n\}} \sum_{s=1}^n |a_{is} - a_{js}| = 1 - \min_{i, j \in \{1, \dots, n\}} \sum_{s=1}^n \min\{a_{is}, a_{js}\}. \quad (\text{A.9})$$

It is known that this coefficient of ergodicity satisfies $0 \leq \tau(\mathbf{A}) \leq 1$, and $\tau(\mathbf{A}) = 0$ if and only if all the rows of \mathbf{A} are identical (i.e. $\mathbf{A} = \mathbf{1}_n \mathbf{z}^\top$ for some $\mathbf{z} \in \mathbb{R}^n$). Importantly, there holds

$$\tau(\mathbf{A}) < 1 \quad (\text{A.10})$$

if $\mathbf{A} > 0$. Last, note that there holds (see e.g. [Seneta, 2006].)

$$V(\mathbf{Ax}) \leq \tau(\mathbf{A})V(\mathbf{x}). \quad (\text{A.11})$$

A.2 Graph Theory

In this section, some additional results are given on graph theory, primarily for use in Chapter 9, where the model considered is continuous-time as opposed to the other chapters, which all consider discrete-time models. Generally speaking,

the graph topology considered by continuous-time models requires the use of the *Laplacian* matrix associated with a graph \mathcal{G} .

Consider a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$. Each individual is a node in the finite, nonempty set of nodes $\mathcal{V} = \{v_i : i \in \mathcal{I} = \{1, \dots, n\}\}$. The set of ordered edges is $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. An ordered edge is denoted as $e_{ij} = (v_i, v_j) \in \mathcal{E}$. An edge e_{ij} is said to be outgoing with respect to v_i and incoming with respect to v_j . The presence of an edge e_{ij} connotes that individual j learns of, and takes into account, the opinion value of individual i when updating its own opinion. The (incoming) neighbour set of v_i is defined as $\mathcal{N}_i = \{v_j \in \mathcal{V} : e_{ji} \in \mathcal{E}\}$. The weighted adjacency matrix $\mathcal{A} \in \mathbb{R}^{n \times n}$ of \mathcal{G} has nonnegative elements a_{ij} . The elements of \mathcal{A} are such that $a_{ij} > 0 \Leftrightarrow e_{ji} \in \mathcal{E}$ while $a_{ij} = 0$ if $e_{ji} \notin \mathcal{E}$ and it is assumed that $a_{ii} = 0, \forall i$. The $n \times n$ Laplacian matrix, $\mathcal{L} = \{l_{ij}\}$, of the associated graph \mathcal{G} is defined

$$l_{ij} = \begin{cases} \sum_{k=1, k \neq i}^n a_{ik} & \text{for } j = i \\ -a_{ij} & \text{for } j \neq i \end{cases}$$

A directed spanning tree is a directed graph formed by directed edges of the graph that connects all the nodes, and where every vertex apart from the root has exactly one parent. A graph is said to contain a directed spanning tree if a subset of the edges forms a directed spanning tree¹. A graph is strongly connected if and only if, for every node v_i , there exists a directed path to every other node v_j . The following is a standard result that will be used throughout this paper.

Lemma A.8 ([Mesbahi and Egerstedt, 2010b]). *Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$ be given. The Laplacian \mathcal{L} associated with \mathcal{G} has a single eigenvalue at 0 if and only if \mathcal{G} has a directed spanning tree. Associated with the single eigenvalue at 0 are the right eigenvector $\mathbf{1}_n$ and left eigenvector $\gamma \geq 0$, with normalisation $\gamma^\top \mathbf{1}_n = 1$. All other eigenvalues have strictly positive real part.*

Note that if the graph \mathcal{G} is undirected, i.e. $\mathcal{A} = \mathcal{A}^\top$ and $\mathcal{L} = \mathcal{L}^\top$, then a consequence of Lemma A.8 is that the eigenvalues of \mathcal{L} are real, nonnegative, and can be ordered as $0 = \lambda_1(\mathcal{L}) < \lambda_2(\mathcal{L}) \leq \dots \leq \lambda_n(\mathcal{L})$.

If the graph contains a directed spanning tree, then there exists an $r \leq n$ such that the reordered nodes v_1, \dots, v_r induce a maximally closed and strongly connected subgraph \mathcal{G}_L . By closed, it is meant that no edges are incoming to \mathcal{G}_L . Denote by \mathcal{G}_F the subgraph induced by the set of nodes v_{r+1}, \dots, v_n . With the nodes reordered, the Laplacian matrix \mathcal{L} associated with \mathcal{G} is expressed as

$$\mathcal{L} = \begin{bmatrix} \mathcal{L}_{11} & \mathbf{0}_{r \times (n-r)} \\ \mathcal{L}_{21} & \mathcal{L}_{22} \end{bmatrix} \quad (\text{A.12})$$

where $\mathcal{L}_{11} \in \mathbb{R}^{r \times r}$ is irreducible. In some literature, this is referred to as the Frobenius normal form of the Laplacian matrix [Li et al., 2010]. If $r = n$ then \mathcal{G} is strongly

¹Some literature use other terms, e.g. rooted out-branching [Mesbahi and Egerstedt, 2010b] or directed rooted tree.

connected, and $\mathcal{L}_{22} \in \mathbb{R}^{0 \times 0}$. If $r < n$ then the matrix $\mathcal{L}_{22} \in \mathbb{R}^{(n-r) \times (n-r)}$ is nonsingular, and its eigenvalues have strictly positive real part [Xia and Cao, 2017]. In addition, there is a right eigenvector associated with $\lambda_1(\mathcal{L}) = 0$, with nonnegative entries, given by $\gamma^\top = [\gamma_1, \dots, \gamma_n]$ with $\gamma_i > 0, i \in \{1, \dots, r\}$. In this thesis, it is assumed that γ is normalised such that $\sum_{i=1}^n \gamma_i = 1$ holds.

A.3 Nonlinear Contraction Analysis

In this section of the appendix, some results on nonlinear contraction analysis of discrete-time systems are presented. These results will be used in Chapters 5 and 6. Nonlinear contraction analysis was first exposed to the control community as a method to study dynamical systems in the now classic [Lohmiller and Slotine, 1998, Section 5], though these results have existed in the wider mathematical community for several decades prior.

Consider a deterministic discrete-time system of the form

$$\mathbf{x}(k+1) = \mathbf{f}_{\sigma(k)}(\mathbf{x}(k), k), \quad (\text{A.13})$$

with $n \times 1$ state vector \mathbf{x} , switching signal $\sigma(k)$, and $n \times 1$ vector-valued function $\mathbf{f}_{\sigma(k)}(\mathbf{x}(k), k)$. It is assumed that \mathbf{f} is smooth, by which it is meant that any required derivative or partial derivative exists, and is continuous. The associated virtual² dynamics is

$$\delta \mathbf{x}(k+1) = \frac{\partial \mathbf{f}_{\sigma(k)}}{\partial \mathbf{x}(k)} \delta \mathbf{x}(k).$$

See Fig. A.1 for an illustration of the virtual dynamics. In a region of interest of the state-space $\mathcal{D} \subseteq \mathbb{R}^n$, define the transformation $\delta \mathbf{z}(k) = \mathbf{\Theta}_{\sigma(k)}(\mathbf{x}(k), k) \delta \mathbf{x}(k)$ where $\mathbf{\Theta}_{\sigma(k)}(\mathbf{x}(k), k) \in \mathbb{R}^{n \times n}$ is uniformly nonsingular. More specifically, uniform nonsingularity means that there exist a real number $\kappa > 0$ and a matrix norm $\|\cdot\|'$ such that $\kappa < \|\mathbf{\Theta}_{\sigma(k)}(\mathbf{x}(k), k)\|' < \kappa^{-1}$ holds for all $\mathbf{x} \in \mathcal{D}$ and $k \geq 0$. If the uniformly nonsingular condition holds, then exponential convergence of $\delta \mathbf{z}$ to $\mathbf{0}_n$ implies, and is implied by, exponential convergence of $\delta \mathbf{x}$ to $\mathbf{0}_n$. The transformed virtual dynamics can be computed as

$$\delta \mathbf{z}(k+1) = \mathbf{F}(k) \delta \mathbf{z}(k), \quad (\text{A.14})$$

where $\mathbf{F}(k) = \mathbf{\Theta}_{\sigma(k+1)}(\mathbf{x}(k+1), k+1) \frac{\partial \mathbf{f}_{\sigma(k)}}{\partial \mathbf{x}(k)} \mathbf{\Theta}_{\sigma(k)}(\mathbf{x}(k), k)^{-1}$ is the transformed Jacobian.

Definition A.1 (Generalised Contraction Region). *Given the discrete-time system in Eq. (A.13), a region of the state space $\mathcal{D} \subseteq \mathbb{R}^n$ is called a generalised contraction region with respect to the metric $\|\delta \mathbf{x}\|_{\mathbf{\Theta}(k),1} = \|\mathbf{\Theta}_{\sigma(k)}(\mathbf{x}(k), k) \delta \mathbf{x}\|_1$ if in that region, $\|\mathbf{F}(k)\|_1 < 1 - \eta$ holds for all k , where $\eta > 0$ is an arbitrarily small constant.*

²The term “virtual” is taken from Lohmiller and Slotine [1998]; $\delta \mathbf{x}$ is a virtual, i.e. infinitesimal, displacement.

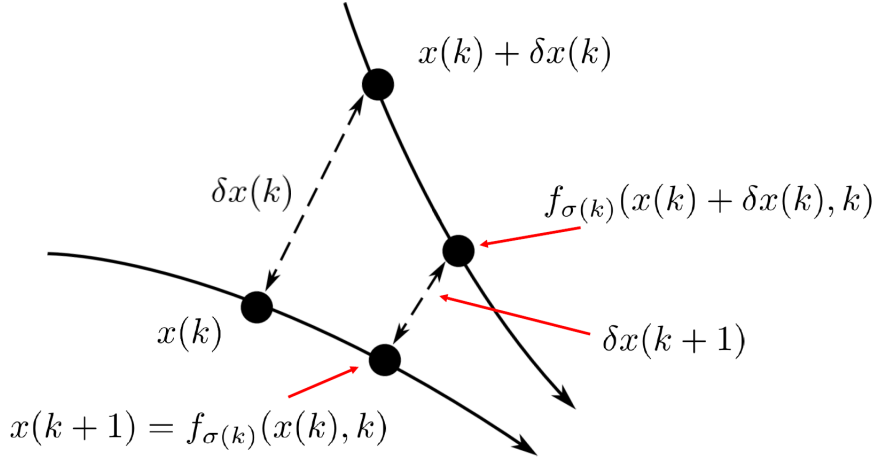


Figure A.1: Two trajectories of $x(k+1) = f_{\sigma(k)}(x(k), k)$ are separated by a virtual displacement $\delta x(k)$. The virtual dynamics are illustrated, and nonlinear contraction analysis is focused on study of how the virtual displacement evolves over k . One desirable property is for the virtual displacement $\delta x(k)$ to converge to zero; this implies certain properties of the system A.13. See Chapters 5 and 6 for the interpretations of these properties in the context of the evolution of social power in social networks.

Note that the 1-norm metric is being considered in the variable space δz which in turn leads to a *weighted* 1-norm in the variable space δx . Here, the weighting matrix is $\Theta_{\sigma(k)}(x(k), k)$ and the weighted 1-norm is well defined over \mathcal{D} because Θ is uniformly nonsingular.

Theorem A.4. *Given the system Eq. (A.13), consider a tube of constant radius with respect to the metric $\|x\|_{\Theta(k),1}$, centred at a given trajectory of Eq. (A.13). Any trajectory, which starts in this tube and is contained at all times in a generalised contraction region, remains in that tube and converges exponentially fast to the given trajectory as $k \rightarrow \infty$.*

Furthermore, global exponential convergence to the given trajectory is guaranteed if the whole state space is a generalised contraction region with respect to the metric $\|x\|_{\Theta(k),1}$.

The now classic paper [Lohmiller and Slotine, 1998] focused on contraction in the Euclidean metric $\|x\|_{\Theta(k),2} = \|\Theta_{\sigma(k)}(x(k), k)x(k)\|_2$. However, it is obvious from Eq. (A.14) that other norms can be considered. This thesis, and specifically Chapters 5 and 6, requires use of the 1-norm metric because the 2-norm metric does not deliver a convergence result (see the relevant chapters for comments and remarks on this issue). A sketch of the proof is provided here, modified for the 1-norm metric, and the reader is referred to [Lohmiller and Slotine, 1998] for precise details.

Proof. In a generalised contraction region, there holds

$$\begin{aligned} \|\delta z(k+1)\|_1 &= \|F(k)\delta z(k)\|_1 \\ \|\delta z(k+1)\|_1 &< (1 - \eta)\|\delta z(k)\|_1 \end{aligned}$$

since $\|F(k)\|_1 < 1 - \eta$ holds $\forall k$ inside the generalised contraction region³. This implies that $\lim_{k \rightarrow \infty} \delta z(k) = \mathbf{0}_n$ exponentially fast, which in turn implies that there holds $\lim_{k \rightarrow \infty} \delta x(k) = \mathbf{0}_n$ exponentially fast due to uniform nonsingularity of $\Theta_{\sigma(k)}(x(k), k)$. The definition of δx then implies that *any two infinitesimally close trajectories* of the system in Eq. (A.13) converge to each other exponentially fast.

The distance between two points, P_1 and P_2 , with respect to the metric⁴ $\|\cdot\|_{\Theta(k),1}$ is defined as the shortest path length between P_1 and P_2 , i.e., the smallest path integral $\int_{P_1}^{P_2} \|\delta z\|_1 = \int_{P_1}^{P_2} \|\delta x\|_{\Theta(k),1}$. A tube centred about a trajectory $x_1(k)$ and with radius R is then defined as the set of all points whose distances to $x_1(k)$ with respect to $\|\cdot\|_{\Theta(k),1}$ are strictly less than R .

Let $x_2(k) \neq x_1(k)$ be any trajectory that starts inside this tube, separated from $x_1(k)$ by a finite distance with respect to the metric $\|\cdot\|_{\Theta(k),1}$. Suppose that the tube is contained at all times in a generalised contraction region. The fact that there holds $\lim_{k \rightarrow \infty} \|\delta x(k)\|_{\Theta(k),1} = 0$ then implies that $\lim_{k \rightarrow \infty} \int_{x_1(k)}^{x_2(k)} \|\delta x(k)\|_{\Theta(k),1} = 0$ exponentially fast. That is, given the trajectories $x_2(k)$ and $x_1(k)$, separated by a finite distance with respect to the metric $\|\cdot\|_{\Theta(k),1}$, $x_2(k)$ converges to $x_1(k)$ exponentially fast. Global convergence is obtained by setting $R = \infty$. \square

Corollary A.1. *If the contraction region is convex, then all trajectories converge exponentially fast to a unique trajectory.*

Proof. This immediately follows from the fact that any finite distance between two trajectories shrinks exponentially in the convex region. \square

³One needs $\eta > 0$ to eliminate the possibility that $\lim_{k \rightarrow \infty} \|F(k)\|_1 = 1$, which would not result in exponential convergence.

⁴It is easy to verify that the requirements of a metric are satisfied.

Bibliography

- Australian Bureau of Statistics Marriage Law Postal Survey. <http://www.abs.gov.au/ausstats/abs@.nsf/mf/1800.0>. Accessed: 2018-02-14. (cited on page 64)
- ABBINK, K.; GANGADHARAN, L.; HANDFIELD, T.; AND THRASHER, J., 2017. Peer punishment promotes enforcement of bad social norms. *Nature Communications*, 8 (2017). doi:10.1038/s41467-017-00731-0. (cited on page 22)
- ABELSON, R. P., 1964. Mathematical models of the distribution of attitudes under controversy. *Contributions to mathematical psychology*, 14 (1964), 1–160. (cited on pages 3 and 152)
- ALTAFINI, C., 2013. Consensus Problems on Networks with Antagonistic Interactions. *IEEE Transactions on Automatic Control*, 58, 4 (2013), 935–946. (cited on pages 4 and 152)
- AMELKIN, V.; BULLO, F.; AND SINGH, A. K., 2017. Polar opinion dynamics in social networks. *IEEE Transactions on Automatic Control*, 62, 11 (2017), 5650–5665. (cited on pages 100 and 152)
- AMELKIN, V. AND SINGH, A. K., 2017. Disabling External Influence in Social Networks via Edge Recommendation. *arXiv preprint arXiv:1709.08139*, (2017). (cited on page 118)
- ANDERSON, B. D. O.; SHI, G.; AND TRUMPF, J., 2017. Convergence and State Reconstruction of Time-Varying Multi-Agent Systems From Complete Observability Theory. *IEEE Transactions on Automatic Control*, 62, 5 (2017), 2519–2523. (cited on page 100)
- ANDERSON, B. D. O. AND YE, M., 2018. Nonlinear Mapping Convergence and Application to Social Networks. In *to appear in European Control Conference*. <https://arxiv.org/abs/1709.08840>. ArXiv:1709.08840 [cs.SY]. (cited on pages 138, 141, and 148)
- ANDERSON, B. D. O.; YU, C.; FIDAN, B.; AND HENDRICKX, J. M., 2008. Rigid graph control architectures for autonomous formations. *Control Systems, IEEE*, 28, 6 (2008), 48–63. (cited on page 8)
- ARMSTRONG, M. A., 2013. *Basic Topology*. Springer Science & Business Media. (cited on pages 139 and 140)

-
- ASCH, S. E., 1951. *Groups, Leadership, and Men*, chap. Effects of Group Pressure Upon the Modification and Distortion of Judgments, 222–236. Carnegie Press: Pittsburgh. (cited on pages 5, 22, 26, 52, 54, 55, 57, 58, 62, and 181)
- BECKER, J.; BRACKBILL, D.; AND CENTOLA, D., 2017. Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences*, 114, 26 (2017), E5070–E5076. doi:10.1073/pnas.1615978114. <http://www.pnas.org/content/114/26/E5070.abstract>. (cited on pages 3, 4, and 87)
- BERMAN, A. AND PLEMMONS, R. J., 1979. Nonnegative Matrices. *The Mathematical Sciences, Classics in Applied Mathematics*, 9 (1979). (cited on page 185)
- BERNSTEIN, D. S., 2009. *Matrix Mathematics: Theory, Facts, and Formulas*. Princeton University Press. (cited on pages 43, 44, 185, and 186)
- BIKHCHANDANI, S.; HIRSHLEIFER, D.; AND WELCH, I., 1992. A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy*, 100, 5 (1992), 992–1026. (cited on page 22)
- BLONDEL, V. D.; HENDRICKX, J. M.; AND TSITSIKLIS, J. N., 2009. On Krause’s Multi-Agent Consensus Model with State-Dependent Connectivity. *IEEE Transactions on Automatic Control*, 54, 11 (2009), 2586–2597. (cited on page 3)
- BLONDEL, V. D. AND TSITSIKLIS, J. N., 2000. The boundedness of all products of a pair of matrices is undecidable. *Systems & Control Letters*, 41, 2 (2000), 135–140. (cited on page 94)
- BOYD, S.; GHOSH, A.; PRABHAKAR, B.; AND SHAH, D., 2006. Randomized gossip algorithms. *IEEE Transactions Information Theory*, 52, 6 (June 2006), 2508–2530. doi:10.1109/TIT.2006.874516. (cited on page 3)
- BRIN, S. AND PAGE, L., 1998. The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems*, 30, 1-7 (1998), 107–117. (cited on page 85)
- BROCKMANN, D. AND HELBING, D., 2013. The Hidden Geometry of Complex, Network-Driven Contagion Phenomena. *Science*, 342, 6164 (2013), 1337–1342. (cited on page 1)
- BULLO, F.; CORTES, J.; AND MARTINEZ, S., 2009. *Distributed Control of Robotic Networks*. Princeton University Press. (cited on pages 12, 13, 14, and 160)
- CAI, K. AND ISHII, H., 2012. Average consensus on general strongly connected digraphs. *Automatica*, 48, 11 (2012), 2750–2761. (cited on page 8)
- CALABRESI, M., 2017. Inside Russia’s Social Media War on America. *Time*, (2017). (cited on page 54)

-
- CAO, M.; MORSE, A. S.; AND ANDERSON, B. D. O., 2008. Reaching a Consensus in a Dynamically Changing Environment: A Graphical Approach. *SIAM Journal on Control and Optimization*, 47, 2 (2008), 575–600. (cited on pages 3, 8, and 100)
- CAO, Y.; REN, W.; AND EGERSTEDT, M., 2012. Distributed containment control with multiple stationary or dynamic leaders in fixed and switching directed networks. *Automatica*, 48, 8 (2012), 1586–1597. (cited on pages 8 and 154)
- CARTWRIGHT, D., 1959. *Studies in Social Power*. Research Center for Group Dynamics, Institute for Social Research. University of Michigan. (cited on page 85)
- CENTOLA, D.; WILLER, R.; AND MACY, M., 2005. The Emperor’s Dilemma: A Computational Model of Self-Enforcing Norms. *American Journal of Sociology*, 110, 4 (2005), 1009–1040. (cited on page 22)
- CHANDRASEKHAR, A. G.; LARREGUY, H.; AND XANDRI, J. P., 2012. Testing models of social learning on networks: Evidence from a framed field experiment. *Work. Pap., Mass. Inst. Technol., Cambridge, MA*, (2012). (cited on page 3)
- CHEN, X.; LIU, J.; BELABBAS, M.-A.; XU, Z.; AND BAŞAR, T., 2017. Distributed Evaluation and Convergence of Self-Appraisals in Social Networks. *IEEE Transactions on Automatic Control*, 62, 1 (2017), 291–304. (cited on pages 85 and 152)
- CHOPRA, N., 2012. Output Synchronization on Strongly Connected Graphs. *IEEE Transactions on Automatic Control*, 57, 11 (Nov 2012), 2896–2901. doi:10.1109/TAC.2012.2193704. (cited on page 8)
- COCH, L. AND FRENCH JR, J. R., 1948. Overcoming Resistance to Change. *Human Relations*, 1, 4 (1948), 512–532. (cited on page 22)
- CONVERSE, P. E., 1964. The nature of belief systems in mass publics. *Ideology and Discontent*, (1964), 206–61. (cited on pages 152 and 156)
- COOLEY, C. H., 1992. *Human Nature and the Social Order*. Transaction Publishers. (cited on page 82)
- COWAN, N. J.; CHASTAIN, E. J.; VILHENA, D. A.; FREUDENBERG, J. S.; AND BERGSTROM, C. T., 2012. Nodal dynamics, Not Degree Distributions, Determine the Structural Controllability of Complex Networks. *PloS one*, 7, 6 (2012), e38398. (cited on page 181)
- CSÁJI, B. C.; JUNGERS, R. M.; AND BLONDEL, V. D., 2014. Pagerank optimization by edge selection. *Discrete Applied Mathematics*, 169 (2014), 73–87. (cited on page 118)
- DANDEKAR, P.; GOEL, A.; AND LEE, D. T., 2013. Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, 110, 15 (2013), 5791–5796. (cited on pages 4, 6, 26, and 100)

-
- DEGROOT, M. H., 1974. Reaching a Consensus. *Journal of the American Statistical Association*, 69, 345 (1974), 118–121. (cited on pages 2, 3, 26, and 87)
- DEL VICARIO, M.; BESSI, A.; ZOLLO, F.; PETRONI, F.; SCALA, A.; CALDARELLI, G.; STANLEY, H. E.; AND QUATTROCIOCCHI, W., 2016. The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113, 3 (2016), 554–559. (cited on pages 1 and 54)
- DUGGINS, P., 2017. A Psychologically-Motivated Model of Opinion Change with Applications to American Politics. *Journal of Artificial Societies and Social Simulation*, 20, 1 (2017), 1–13. doi:10.18564/jasss.3316. <http://jasss.soc.surrey.ac.uk/20/1/13.html>. (cited on pages 4, 7, 26, 36, and 100)
- ETESAMI, S. R. AND BAŞAR, T., 2015. Game-Theoretic Analysis of the Hegselmann-Krause Model for Opinion Dynamics in Finite Dimensions. *IEEE Transactions on Automatic Control*, 60, 7 (2015), 1886–1897. (cited on page 3)
- FERCOQ, O.; AKIAN, M.; BOUHTOU, M.; AND GAUBERT, S., 2013. Ergodic Control and Polyhedral Approaches to PageRank Optimization. *IEEE Transactions on Automatic Control*, 58, 1 (2013), 134–148. (cited on page 118)
- FESTINGER, L., 1950. Informal Social Communication. *Psychological Review*, 57, 5 (1950), 271. (cited on page 26)
- FLACHE, A.; MÄS, M.; FELICIANI, T.; CHATTOE-BROWN, E.; DEFFUANT, G.; HUET, S.; AND LORENZ, J., 2017. Models of Social Influence: Towards the Next Frontiers. *Journal of Artificial Societies & Social Simulation*, 20, 4 (2017). (cited on page 2)
- FRENCH JR, J. R. P., 1956. A Formal Theory of Social Power. *Psychological Review*, 63, 3 (1956), 181–194. (cited on pages 2, 82, and 152)
- FRENCH JR, J. R. P. AND SNYDER, R., 1959. Leadership and Interpersonal Power. In *Studies in Social Power* (Ed. D. CARTWRIGHT), chap. 8, 118–149. Research Center for Group Dynamics, Institute for Social Research, University of Michigan. (cited on page 118)
- FRIEDKIN, N. E., 2011. A Formal Theory of Reflected Appraisals in the Evolution of Power. *Administrative Science Quarterly*, 56, 4 (2011), 501–529. (cited on page 118)
- FRIEDKIN, N. E., 2015. The Problem of Social Control and Coordination of Complex Systems in Sociology: A Look at the Community Cleavage Problem. *IEEE Control Systems Magazine*, 35, 3 (2015), 40–51. (cited on page 2)
- FRIEDKIN, N. E. AND BULLO, F., 2017. How truth wins in opinion dynamics along issue sequences. *Proceedings of the National Academy of Sciences*, (2017). doi:10.1073/pnas.1710603114. <http://www.pnas.org/content/early/2017/10/04/1710603114.abstract>. (cited on pages 5, 16, and 181)

-
- FRIEDKIN, N. E.; JIA, P.; AND BULLO, F., 2016a. A Theory of the Evolution of Social Power: Natural Trajectories of Interpersonal Influence Systems along Issue Sequences. *Sociological Science*, 3 (2016), 444–472. (cited on pages 6 and 87)
- FRIEDKIN, N. E. AND JOHNSEN, E. C., 1990. Social Influence and Opinions. *Journal of Mathematical Sociology*, 15, 3-4 (1990), 193–206. (cited on pages 4, 16, 26, and 152)
- FRIEDKIN, N. E. AND JOHNSEN, E. C., 2011. *Social Influence Network Theory: A Sociological Examination of Small Group Dynamics*, vol. 33. Cambridge University Press. (cited on pages 5, 62, and 181)
- FRIEDKIN, N. E.; PROSKURNIKOV, A. V.; TEMPO, R.; AND PARSEGOV, S. E., 2016b. Network science on belief system dynamics under logic constraints. *Science*, 354, 6310 (2016), 321–326. (cited on pages 5, 6, 16, 152, 153, and 156)
- GAWRONSKI, B. AND STRACK, F., 2012. *Cognitive Consistency: A Fundamental Principle in Social Cognition*. Guilford press. (cited on page 156)
- GECAS, V. AND SCHWALBE, M. L., 1983. Beyond the Looking-Glass Self: Social Structure and Efficacy-Based Self-Esteem. *Social Psychology Quarterly*, (1983), 77–88. (cited on page 82)
- GLEICH, D. F., 2015. PageRank Beyond the Web. *SIAM Review*, 57, 3 (2015), 321–363. (cited on page 85)
- GODSIL, C. D. AND ROYLE, G., 2001. *Algebraic Graph Theory*, vol. 207. Springer: New York. (cited on page 13)
- GOODWIN, J., 2011. Why We Were Surprised (Again) by the Arab Spring. *Swiss Political Science Review*, 17, 4 (2011), 452–456. (cited on pages 22 and 36)
- GORDEN, R. L., 1952. Interaction Between Attitude and the Definition of the Situation in the Expression of Opinion. *American Sociological Review*, 17, 1 (1952), 50–58. (cited on pages 22, 26, and 181)
- GUILLEMIN, V. AND POLLACK, A., 2010. *Differential Topology*, vol. 370. American Mathematical Soc. (cited on pages 139, 140, and 141)
- HARARY, F., 1959. *Studies in social power*, chap. A criterion for unanimity in French's theory of social power, 198–182. Ann Arbor, University of Michigan Press. (cited on pages 2 and 3)
- HARARY, F.; NORMAN, R. Z.; AND CARTWRIGHT, D., 1965. *Structural Models: An Introduction to the Theory of Directed Graphs*. J. Wiley & Sons. (cited on pages 2 and 3)
- HEGSELMANN, R. AND KRAUSE, U., 2002. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation*, 5, 3 (2002). (cited on pages 3, 26, and 82)

-
- HENDRICKX, J. M., 2014. A lifting approach to models of opinion dynamics with antagonisms. In *IEEE 53rd Annual Conference on Decision and Control (CDC)*, 2118–2123. IEEE. (cited on page 152)
- HENDRICKX, J. M.; SHI, G.; AND JOHANSSON, K. H., 2015. Finite-Time Consensus Using Stochastic Matrices With Positive Diagonals. *IEEE Transactions on Automatic Control*, 60, 4 (2015), 1070–1073. (cited on pages 15 and 86)
- HIRSCH, M. W., 2012. *Differential Topology*, vol. 33. Springer Science & Business Media. (cited on pages 139, 140, and 141)
- HORN, R. A. AND JOHNSON, C. R., 2012. *Matrix Analysis*. Cambridge University Press, New York. (cited on pages 94, 148, and 186)
- ISHII, H. AND TEMPO, R., 2014. The PageRank Problem, Multiagent Consensus, and Web Aggregation: A Systems and Control Viewpoint. *IEEE Control Systems*, 34, 3 (2014), 34–53. (cited on page 86)
- JADBABAIE, A.; LIN, J.; AND MORSE, A. S., 2003. Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules. *IEEE Transactions on Automatic Control*, 48, 6 (June 2003), 988–1001. doi:10.1109/TAC.2003.812781. (cited on pages 8 and 100)
- JIA, P.; FRIEDKIN, N. E.; AND BULLO, F., 2017a. Opinion Dynamics and Social Power Evolution: A Single-Timescale Model. *arXiv preprint arXiv:1712.02452*, (2017). (cited on page 182)
- JIA, P.; FRIEDKIN, N. E.; AND BULLO, F., 2017b. Opinion Dynamics and Social Power Evolution over Reducible Influence Networks. *SIAM Journal on Control and Optimization*, 55, 2 (2017), 1280–1301. (cited on page 83)
- JIA, P.; MIRTABATABAEI, A.; FRIEDKIN, N. E.; AND BULLO, F., 2015. Opinion Dynamics and the Evolution of Social Power in Influence Networks. *SIAM Review*, 57, 3 (2015), 367–397. (cited on pages 6, 81, 82, 83, 84, 85, 86, 87, 88, 89, 93, 96, 97, 98, 104, 106, 118, 120, 121, 138, 143, and 148)
- KEMPE, D.; KLEINBERG, J.; AND TARDOS, E., 2003. Maximizing the Spread of Influence through a Social Network. In *Proceedings of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 137–146. (cited on pages 1 and 118)
- KHALIL, H., 2002. *Nonlinear Systems*. Prentice Hall. (cited on page 93)
- KHAMSI, M. A. AND KIRK, W. A., 2011. *An Introduction to Metric Spaces and Fixed Point Theory*. John Wiley & Sons. (cited on pages 93, 94, and 148)
- KRICK, L.; BROUCKE, M. E.; AND FRANCIS, B. A., 2009. Stabilisation of infinitesimally rigid formations of multi-robot networks. *International Journal of Control*, 82, 3 (2009), 423–439. (cited on page 8)

-
- KURAN, T., 1989. Sparks and prairie fires: A theory of unanticipated political revolution. *Public Choice*, 61, 1 (1989), 41–74. (cited on pages 22 and 36)
- KURAN, T., 1997. *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Harvard University Press. (cited on pages 5, 24, 53, 78, and 181)
- LI, Z.; DUAN, Z.; CHEN, G.; AND HUANG, L., 2010. Consensus of Multiagent Systems and Synchronization of Complex Networks: A Unified Viewpoint. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 57, 1 (2010), 213–224. (cited on page 188)
- LIN, C.-T., 1974. Structural Controllability. *IEEE Transactions on Automatic Control*, 19, 3 (1974), 201–208. (cited on page 181)
- LIU, B. J.; MOU, S.; MORSE, A. S.; ANDERSON, B.; AND YU, C., 2011. Deterministic gossiping. *Proceedings of the IEEE*, 99, 9 (2011), 1505–1524. (cited on page 3)
- LIU, F. AND MORSE, A. S., 2017. Structural Controllability of Linear Systems. In *Decision and Control (CDC), 2017 IEEE 56th Annual Conference on*, 3588–3593. IEEE. (cited on page 181)
- LIU, J.; CHEN, X.; BAŞAR, T.; AND BELABBAS, M.-A., 2017a. Exponential Convergence of the Discrete- and Continuous-Time Altafini Models. *IEEE Transaction on Automatic Control*, 62, 12 (2017), 6168–6182. (cited on pages 4, 100, and 152)
- LIU, J.; YE, M.; ANDERSON, B. D. O.; BAŞAR, T.; AND NEDIĆ, A., 2017b. Discrete-Time Polar Opinion Dynamics. <https://arxiv.org/abs/1709.07121>. ArXiv:1709.07121 [cs.SI]. (cited on page 152)
- LOHMILLER, W. AND SLOTINE, J.-J. E., 1998. On contraction analysis for non-linear systems. *Automatica*, 34, 6 (1998), 683–696. (cited on pages 94, 189, and 190)
- LORENZ, J., 2007. Continuous Opinion Dynamics Under Bounded Confidence: A Survey. *International Journal of Modern Physics C*, 18, 12 (2007), 1819–1838. (cited on page 3)
- MÄS, M.; FLACHE, A.; AND KITTS, J. A., 2014. Cultural Integration and Differentiation in Groups and Organizations. In *Perspectives on Culture and Agent-based Simulations*, 71–90. Springer. (cited on pages 4, 7, 26, and 100)
- MATSUMOTO, Y., 2002. *An Introduction to Morse Theory*, vol. 208. American Mathematical Soc. (cited on pages 140 and 141)
- MCCRIGHT, A. M. AND DUNLAP, R. E., 2017. Combatting Misinformation Requires Recognizing Its Types and the Factors That Facilitate Its Spread and Resonance. vol. 6, 389–396. Elsevier. (cited on pages 54, 63, and 78)
- MEI, W.; MOHAGHEGHI, S.; ZAMPIERI, S.; AND BULLO, F., 2017. On the dynamics of deterministic epidemic propagation over networks. *Annual Reviews in Control*, (2017). (cited on page 1)

- MELLON, J.; YODER, J.; AND EVANS, D., 2016. Undermining and Strengthening Social Networks Through Network Modification. *Scientific Reports*, 6 (2016), 34613. (cited on pages 1 and 118)
- MEREI, F., 1949. Group Leadership and Institutionalization. *Human Relations*, 2, 1 (1949), 23–39. (cited on page 26)
- MERTON, R. K., 1968. *Social Theory and Social Structure*. Free Press: New York. (cited on page 63)
- MESBAHI, M. AND EGERSTEDT, M., 2010a. *Graph Theoretic Methods in Multiagent Networks*. Princeton University Press, Princeton. (cited on page 100)
- MESBAHI, M. AND EGERSTEDT, M., 2010b. *Graph theoretic methods in multiagent networks*. Princeton University Press. (cited on pages 154 and 188)
- MINTZ, A. AND WAYNE, C., 2016. *The Polythink Syndrome: US Foreign Policy Decisions on 9/11, Afghanistan, Iraq, Iran, Syria, and ISIS*. Stanford University Press. (cited on page 22)
- MIRTABATABAEI, A. AND BULLO, F., 2012. Opinion Dynamics in Heterogeneous Networks: Convergence Conjectures and Theorems. *SIAM Journal on Control and Optimization*, 50, 5 (2012), 2763–2785. (cited on page 3)
- MIRTABATABAEI, A.; JIA, P.; FRIEDKIN, N. E.; AND BULLO, F., 2014. On the reflected appraisals dynamics of influence networks with stubborn agents. In *2014 American Control Conference*, 3978–3983. doi:10.1109/ACC.2014.6859256. (cited on page 87)
- NEDIĆ, A. AND LIU, J., 2017. On Convergence Rate of Weighted-Averaging Dynamics for Consensus Problems. *IEEE Transactions on Automatic Control*, 62, 2 (2017), 766–781. (cited on pages 3, 14, and 100)
- NEWMAN, M., 2010. *Networks: An Introduction*. Oxford University Press. (cited on pages 13, 64, 65, 66, 78, and 85)
- NOELLE-NEUMANN, E., 1993. *The Spiral of Silence: Public Opinion, Our Social Skin*. University of Chicago Press. (cited on pages 52 and 183)
- NOWAK, A.; SZAMREJ, J.; AND LATANÉ, B., 1990. From Private Attitude to Public Opinion: A Dynamic Theory of Social Impact. *Psychological Review*, 97, 3 (1990), 362. (cited on pages 17 and 26)
- O’GORMAN, H. J., 1975. Pluralistic Ignorance and White Estimates of White Support for Racial Segregation. *Public Opinion Quarterly*, 39, 3 (1975), 313–330. (cited on pages 5, 52, and 63)
- OLFATI-SABER, R.; FAX, J.; AND MURRAY, R., 2007. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95, 1 (Jan 2007), 215–233. doi:10.1109/JPROC.2006.887293. (cited on page 8)

-
- PARSEGOV, S. E.; PROSKURNIKOV, A. V.; TEMPO, R.; AND FRIEDKIN, N. E., 2017. Novel Multidimensional Models of Opinion Dynamics in Social Networks. *IEEE Transactions on Automatic Control*, 62, 5 (2017), 2270–2285. (cited on pages 5, 6, 16, 152, 153, 156, 157, 159, 169, and 174)
- PRECIADO, V. M.; ZARGHAM, M.; ENYIOHA, C.; JADBABAIE, A.; AND PAPPAS, G., 2014. Optimal Resource Allocation for Network Protection: A Geometric Programming Approach. *IEEE Transactions on Control of Network Systems*, 1, 1 (2014), 99–108. (cited on page 118)
- PROSKURNIKOV, A.; MATVEEV, A.; AND CAO, M., 2016. Opinion dynamics in social networks with hostile camps: Consensus vs. polarization. *IEEE Transaction on Automatic Control*, 61, 6 (2016), 1524–1536. (cited on pages 4, 100, and 152)
- PROSKURNIKOV, A. V. AND TEMPO, R., 2017. A tutorial on modeling and analysis of dynamic social networks. Part I. *Annual Reviews in Control*, 43 (2017), 65–79. (cited on pages 2, 7, 14, 15, 86, and 154)
- PROSKURNIKOV, A. V.; TEMPO, R.; CAO, M.; AND FRIEDKIN, N. E., 2017. Opinion evolution in time-varying social influence networks with prejudiced agents. In *IFAC-PapersOnLine*, vol. 50, 11896–11901. Elsevier. (cited on page 17)
- QIN, J. AND YU, C., 2013. Cluster consensus control of generic linear multi-agent systems under directed topology with acyclic partition. *Automatica*, 49, 9 (2013), 2898–2905. (cited on page 8)
- QIN, J.; YU, C.; AND GAO, H., 2015. Collective Behavior for Group of Generic Linear Agents Interacting under Arbitrary Network Topology. *IEEE Transactions on Control of Network Systems*, 2, 3 (2015), 288–297. (cited on page 155)
- QU, Z., 2009. *Cooperative Control of Dynamical Systems: Applications to Autonomous Vehicles*. Springer Science & Business Media. (cited on page 169)
- RAMAZI, P.; RIEHL, J.; AND CAO, M., 2016. Networks of conforming or nonconforming individuals tend to reach satisfactory decisions. *Proceedings of the National Academy of Sciences*, 113, 46 (2016), 12985–12990. (cited on page 1)
- RAVAZZI, C.; FRASCA, P.; TEMPO, R.; AND ISHII, H., 2015. Ergodic Randomized Algorithms and Dynamics Over Networks. *IEEE Transactions on Control of Network Systems*, 2, 1 (2015), 78–87. (cited on page 3)
- REN, W. AND BEARD, R., 2007. *Distributed Consensus in Multi-vehicle Cooperative Control: Theory and Applications*. Springer London. (cited on pages 14 and 100)
- REN, W. AND BEARD, R. W., 2005. Consensus Seeking in Multiagent Systems Under Dynamically Changing Interaction Topologies. *IEEE Transactions on Automatic Control*, 50, 5 (2005), 655–661. (cited on pages 3, 8, and 100)

-
- REN, W. AND CAO, Y., 2011. *Distributed Coordination of Multi-agent Networks: Emergent Problems, Models and Issues*. Springer London. (cited on page 158)
- RIZOIU, M.-A.; GRAHAM, T.; ZHANG, R.; ZHANG, Y.; ACKLAND, R.; AND XIE, L., 2018. # debatenight: The role and influence of socialbots on twitter during the 1st us presidential debate. In *International AAAI Conference on Web and Social Media*, arXiv preprint arXiv:1802.09808. (cited on page 54)
- RUGH, W. J., 1996. *Linear System Theory*. Prentice Hall, Upper Saddle River, New Jersey, second edn. (cited on page 31)
- RYAN, B. AND GROSS, N. C., 1943. The diffusion of hybrid seed corn in two Iowa communities. *Rural Sociology*, 8, 1 (1943), 15. (cited on page 22)
- SCHACHTER, S., 1951. Deviation, Rejection, and Communication. *The Journal of Abnormal and Social Psychology*, 46, 2 (1951), 190–207. (cited on pages 22, 26, and 181)
- SENETA, E., 2006. *Non-negative Matrices and Markov Chains*. Springer Science & Business Media. (cited on page 187)
- SHI, G. AND JOHANSSON, K. H., 2013. The Role of Persistent Graphs in the Agreement Seeking of Social Networks. *IEEE Journal on Selected Areas in Communications*, 31, 9 (September 2013), 595–606. doi:10.1109/JSAC.2013.SUP.0513052. (cited on pages 3 and 100)
- SHRAUGER, J. S. AND SCHOENEMAN, T. J., 1979. Symbolic interactionist view of self-concept: Through the looking glass darkly. *Psychological Bulletin*, 86, 3 (1979), 549. (cited on page 82)
- SMALDINO, P. E. AND EPSTEIN, J. M., 2015. Social conformity despite individual preferences for distinctiveness. *Royal Society Open Science*, 2, 3 (2015), 140437. (cited on pages 4, 7, and 78)
- SPONG, M. W.; HUTCHINSON, S.; AND VIDYASAGAR, M., 2006. *Robot Modeling and Control*, vol. 3. Wiley New York. (cited on page 6)
- SU, W.; CHEN, G.; AND HONG, Y., 2017. Noise leads to quasi-consensus of hegselmann–krause opinion dynamics. *Automatica*, 85 (2017), 448 – 454. doi:https://doi.org/10.1016/j.automatica.2017.08.008. (cited on page 3)
- TAYLOR, D. G., 1982. Pluralistic Ignorance and the Spiral of Silence: A Formal Analysis. *Public Opinion Quarterly*, 46, 3 (1982), 311–335. (cited on pages 53 and 183)
- TAYLOR, M., 1968. Towards a Mathematical Theory of Influence and Attitude Change. *Human Relations*, 21, 2 (1968), 121–139. (cited on pages 5, 152, 153, 154, and 155)
- THRASHER, F. M., 1963. *The Gang: A Study of 1,313 Gangs in Chicago*. University of Chicago Press. (cited on page 22)

-
- TSITSIKLIS, J. N. AND BLONDEL, V. D., 1997. The Lyapunov Exponent and Joint Spectral Radius of Pairs of Matrices Are Hard—When Not Impossible—To Compute and To Approximate. *Mathematics of Control, Signals, and Systems*, 10, 1 (1997), 31–40. (cited on page 94)
- VARGA, R. S., 2009. *Matrix Iterative Analysis*, vol. 27. Springer Science & Business Media. (cited on page 185)
- WANG, J.; YU, X.; AND STONE, L., 2016. Effective Augmentation of Complex Networks. *Scientific reports*, 6 (2016), 25627. (cited on page 118)
- WATERS, N. L. AND HANS, V. P., 2009. A Jury of One: Opinion Formation, Conformity, and Dissent on Juries. *Journal of Empirical Legal Studies*, 6, 3 (2009), 513–540. (cited on pages 5, 22, 180, and 181)
- WILLER, R.; KUWABARA, K.; AND MACY, M. W., 2009. The False Enforcement of Unpopular Norms. *American Journal of Sociology*, 115, 2 (2009), 451–490. (cited on page 22)
- WOOD, W.; LUNDGREN, S.; OUELLETTE, J. A.; BUSCEME, S.; AND BLACKSTONE, T., 1994. Minority influence: A meta-analytic review of social influence processes. (cited on page 78)
- WU, Y.; SU, H.; SHI, P.; LU, R.; AND WU, Z. G., 2016. Output Synchronization of Nonidentical Linear Multiagent Systems. *IEEE Transactions on Cybernetics*, PP, 99 (2016), 1–12. doi:10.1109/TCYB.2015.2508604. (cited on page 8)
- XIA, W. AND CAO, M., 2017. Analysis and applications of spectral properties of grounded laplacian matrices for directed networks. *Automatica*, 80 (2017), 10–16. (cited on page 189)
- XIA, W.; CAO, M.; AND JOHANSSON, K., 2016a. Structural Balance and Opinion Separation in Trust–Mistrust Social Networks. *IEEE Transactions on Control of Network Systems*, 3, 1 (2016), 46–56. (cited on page 152)
- XIA, W.; LIU, J.; JOHANSSON, K. H.; AND BAŞAR, T., 2016b. Convergence Rate of the Modified DeGroot-Friedkin Model with Doubly Stochastic Relative Interaction Matrices. In *American Control Conference (ACC), Boston, USA*, 1054–1059. doi:10.1109/ACC.2016.7525054. (cited on pages 85, 86, and 152)
- XU, Y. AND CORBETT, J., 2015. Measuring financial integration: The network approach. <https://dx.doi.org/10.2139/ssrn.2618283>. (cited on page 85)
- XU, Z.; LIU, J.; AND BAŞAR, T., 2015. On a Modified DeGroot-Friedkin Model of Opinion Dynamics. In *American Control Conference (ACC), Chicago, USA*, 1047–1052. doi:10.1109/ACC.2015.7170871. (cited on pages 85 and 152)

- YE, M.; LIU, J.; AND ANDERSON, B. D. O., 2018. On the Effects of Heterogeneous Logical Interdependencies in Multi-Dimensional Opinion Dynamics Models. In to appear in *IEEE 57th Annual Conference on Decision and Control*. (cited on page 182)
- YEUNG, K.-T. AND MARTIN, J. L., 2003. The Looking Glass Self: An Empirical Test and Elaboration. *Social Forces*, 81, 3 (2003), 843–879. (cited on page 82)
- YILDIZ, E.; OZDAGLAR, A.; ACEMOGLU, D.; SABERI, A.; AND SCAGLIONE, A., 2013. Binary Opinion Dynamics with Stubborn Agents. *ACM Transactions on Economics and Computation*, 1, 4 (2013), 19. (cited on pages 17 and 78)
- YOU, K.; TEMPO, R.; AND QIU, L., 2017. Distributed Algorithms for Computation of Centrality Measures in Complex Networks. *IEEE Transactions on Automatic Control*, 62, 5 (2017), 2080–2094. (cited on page 86)
- YU, C.; ANDERSON, B. D. O.; MOU, S.; LIU, J.; HE, F.; AND MORSE, A. S., 2017. Distributed Averaging Using Periodic Gossiping. *IEEE Transactions on Automatic Control*, 62, 8 (2017), 4282–4289. (cited on page 3)
- ZHAI, X.; ZHENG, L.; WANG, J.; AND TAN, C. W., 2013. Optimization Algorithms for Epidemic Evolution in Broadcast Networks. In *IEEE Wireless Communications and Networking Conference*, 1540–1545. IEEE. (cited on page 118)